

THIS WEEK

EDITORIALS

ENGINEERING Scanner takes the temperature of chemical reactions **p.410**

WORLD VIEW Arab education must start a fire in student minds **p.411**



FACE-OFF Texan pumas leave looks of Florida panthers intact **p.412**

High maintenance

The next president of the European Research Council will face the dual challenge of preserving the agency's reputation for excellence while trying to address funding inequalities.

It is an open secret that French mathematician Jean-Pierre Bourguignon, director of the Institute for Advanced Scientific Studies in Paris, heads the shortlist of candidates for the next president of the European Research Council (ERC).

His expertise in differential geometry might not directly help him to handle the delicate, differential politics that are rife in the European Union (EU), and the consequent tensions between richer western and poorer eastern member states that are the most potent threat to the ERC. But his reputation as a strong-minded defender of the value of research excellence surely will. Such strength is needed to maintain the ERC's happy status quo.

The ERC is a resounding success story. Founded just six years ago to fund highly competitive basic research, it launched itself with an appropriately rigorous — some might say remorseless — peer-review system to fund the best scientists through its two main grant streams. Its reputation for scientific excellence was quickly established, with universities using the number of their ERC grant recipients as a measure of their own status. Winning an ERC grant is an occasion for champagne, for both the honour and the cash — grants are worth up to €3.5 million (US\$4.8 million). Moreover, the ERC is likely to enjoy a significant hike in budget in the European Commission's seven-year Horizon 2020 research-funding programme, which launches in January.

The clouds that threaten this sunny landscape are distant. But they are there, and the challenge will be to keep them at bay. The problem of the gap between rich and poor countries will not disappear any time soon. And, not unexpectedly, such inequality is writ large in ERC statistics. At one extreme, almost half of all ERC grants are awarded to scientists in just three countries: the United Kingdom, Germany and France. At the other, barely 2% are awarded in the former communist countries that joined the EU after 2004.

When politicians in eastern Europe look at these statistics, they are rightly indignant — but they are wrong to ask the ERC to change. They often argue inappropriately for reduced investment in the elite research agency, or for a special ERC funding stream to favour their own disadvantaged countries. The appropriate response would be to fix the problems at home that make their scientists relatively uncompetitive. The countries need to make good use of generous EU structural funds to improve their national research infrastructure. And they need to be more wholehearted in following the EU spirit, laid out in various treaties and agreements, of investing more in national science and allocating most research money competitively. Once their scientists are better placed to compete for ERC grants, the differential will slowly be reduced. But politics is notoriously impatient, and accusations of political discrimination can be powerful.

The commission is unlikely to reveal the new president's identity formally until Horizon 2020 — currently stalled in tense negotiations about the overall EU budget — is signed off towards the end of the year.

The ERC is independent, but needs a strong leader to keep it out

of the sphere of political influence. That is because it is funded by the European Commission, the policies of which are dictated by its political masters, the European Parliament and the European Council. The more beloved and successful the ERC becomes — as indicated by the likely rise in its budget from €7.5 billion now to nearly €12 billion next year, or around 17% of the proposed total Horizon 2020 budget — the more politicians will squabble over who should benefit from its grants.

The ERC is somewhat sheltered from this squabbling because the current leaders of the commission's directorate-general for research and innovation are strong proponents of the ERC. But the leadership will be renewed next October, and the successors might not be so devoted. In any case, the shelter itself can be a double-edged sword. If not kept in check, the commission's byzantine accountability rules would throttle scientists with red tape. The level of detail required for reporting how ERC grant recipients have spent their money is already much too high. The new ERC president will have to ensure that this does not worsen.

The president will also have to maintain attention on the ERC gender gap. According to the latest statistics, only 25% of its grant applicants are women, and their overall success rate is just 8%, compared with 11% for men. The ERC takes many soft measures to try to improve this, mostly through information campaigns, and this needs to continue. And although the ERC budget has improved gratifyingly, with success rates for grant applicants hovering around 10%, it is still much too low for its mission. The president will have to lobby for a level of funding that allows this success rate to double. That, unlike his or her official identity, is no secret at all. ■

"The ERC needs a strong leader to keep it out of the sphere of political influence."

End harassment

Sexual harassment is a stain on science — and we must all take a stand against it.

The past week has seen an outpouring of online comment on the subject of sexual harassment in science and its satellite careers such as science journalism and communication. It was prompted by allegations against a leading figure in science blogging, Bora Zivkovic, who has since resigned as blogs editor with *Scientific American* (which is published by Nature Publishing Group).

Much of the comment has been from women, a distressingly large number of whom have described their own experiences of misogyny and prejudice in the workplace. One lesson to be drawn

is that language matters: in effect, there is no such thing as 'casual' or 'low-level' abuse. And, as the ongoing comments from both men and women on social media make clear, the impact of such behaviour on women, many of whom are early in their careers, can be pernicious and long-lasting. Women can begin to doubt their achievements and their abilities. They might question the motives of people who comment on their work. In short, they can lose confidence; when combined with the structural and institutional obstacles that they already face, this can make women look elsewhere for job satisfaction. This is unacceptable. Science simply cannot afford to lose some of its best talent to boorishness.

A major problem is the widespread tacit acceptance of adolescent behaviour. Let us call him Dr Inappropriate: he is the lecturer at the conference drinks reception with the wandering hands. (No such behaviour has been attributed to Zivkovic.) He is the head of department who thanks his female colleague for her excellent presentation but suggests that she wears a shorter skirt next time (yes, this really happened). Worse, Dr Inappropriate is often the lab head, or an equivalent — a mentor with responsibility and power over the careers of the women whom he asks to work late on a project or to join him in a taxi home. Sometimes he is a very senior scientist indeed.

Nature acknowledged in an Editorial last year that we have poor representation of women among reviewers and authors (see *Nature* 491, 495; 2012) — but we pledged to change and have attempted to do so, with mixed results that we shall report soon. We have asked others to acknowledge their own gender biases, and urged them to do what they can to improve the prospects and visibility of women in science.

Our Women in Science special issue this year (see nature.com/women) offered our most comprehensive and high-profile collection of articles on the subject so far. Yet we have not adequately addressed the problem of harassment, perhaps because it is difficult to quantify. Officially, the obstacles to women in science are policy issues such as availability of childcare and lack of flexible hours. We might never know how many are pushed to leave because they are fed up of

working with Dr Inappropriate. Just as worrying are those women who do not make that choice and who find that they must simply endure.

The evidence of the scale and depth of the problem is anecdotal. But the anecdotes all point to sexual harassment being a real stain on science. Just ask around: everyone knows a Dr Inappropriate. (We have here emphasized male–female harassment, but female–male and same-sex harassment happens too.)

“Science simply cannot afford to lose some of its best talent to boorishness.”

What is to be done? Most institutions already have policies that outlaw harassment and bullying. Could and should they be more strictly enforced? Yes. This often requires the victim to make an official complaint, and many are justifiably reluctant to do so, but a facility for anonymous whistle-blowing may help. A more pragmatic solution is to force Dr Inappropriate to keep his hands to himself, and this is where the rest of us can come in. More of us must challenge such actions when we see them, publicly if necessary. Too often we accommodate and excuse them: “He doesn’t mean it”; “That’s what he’s like after a drink”; “Just make sure you don’t work late on your own.”

There are many behaviours that could be construed as abuse, and there are grey zones. Flirting is human nature. Some students marry their supervisors. Such considerations argue against glib judgements, but must not distract from the central message.

Here is one category of sexual harassment to focus on: when it represents an abuse of a professional relationship, particularly one in which the abuser has power and the victim feels unable to challenge it as they would like. That is wrong, and we should all label it so. We should all seek to promote not only appropriate rules, but also a culture of active discouragement and prevention of sexual harassment. If you are the party with the power, ask yourself: will the recipient of your social overtures wonder whether your support for his or her work is dependent on how she or he responds? If the answer is yes — or even maybe — do not cross that line. ■

Magnetic map

Chemists present a way to infer the enigmatic temperature variations inside a reactor.

Most chemical products start their lives as oil. And most of the conversion processes used to turn the black stuff into plastics, fuels and the rest rely on catalysts. Given the sensitivity of catalysts and Earth’s dwindling supplies of oil, you might think that these reactions would be among the most studied and the best understood in the chemist’s cookbook.

Unfortunately not. In fact, for many chemists and chemical engineers — those who work with bucketloads of reactants rather than the contents of pipettes — what goes on inside an industrial reactor is something of a mystery. It’s a black box. Indeed, when some textbooks and academic papers on the subject show flow charts of chemical processes, they actually represent the reactor, the beating heart of our industrial society, as a black box. If process engineers want to know what happens inside — and so how to make it more efficient, safer or more environmentally friendly — they measure what comes out, compare it with what goes in, and make an educated guess.

As computing power has grown, this educated guesswork has been renamed ‘modelling’. Reconstructions of the catalytic processes that occur in reactors use complex mathematics to represent the relationship between reactants, products and everything in between. Heat transfer, fluid dynamics and surface-reaction kinetics all offer a theoretical platform for such models, but, like all models, they rely on

observations from the real world to make them realistic. Which takes us back to the black box and, often, to the most basic of questions — just how hot is it in there?

Anyone who has cooked a soufflé will know that the temperature, and how it fluctuates inside the oven, has a crucial bearing on the result. They know that the temperature selected and that the oven reaches can disagree. And they know that, even with the best temperature circulation, cool spots can lurk between lower shelves or above a baking tray. Now imagine that your precious pudding relies on the random collisions of a fizzing tempest of high-pressure gas and ageing, unpredictable catalysts. And that you are being asked to deliver 3,000 puddings an hour.

A reliable temperature map of the guts of a working chemical reactor would be valuable. People have tried to achieve this, most often by placing sensors at strategic points. The problem is the age-old paradox that the measurement disturbs what is being measured.

On page 537 of this issue, chemists offer a solution. Nanette Jarenwattananon at the University of California, Los Angeles, and her colleagues describe how they use the magnetic field of an nuclear magnetic resonance (NMR) scanner to accurately infer the hot and cold spots of a reactor carrying out the hydrogenation of propylene. And they report that, under the right conditions, hotter parts of the reactor signal narrower peaks on the NMR spectra.

There is a pleasing symmetry here. In the 1970s, NMR was handed to biologists and renamed magnetic resonance imaging (MRI). The biologists worked out a way to use MRI to sense the temperature inside the human body remotely. Now the chemists have reclaimed both the tool and the function. It is a proof of concept at this stage, but it does go some way towards opening that mysterious black box. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhnuq

A. AWAD



Universities must inspire students as well as teach

Education in the Arab world must equip students with more than textbook learning as they go forward into an uncertain future, says Rana Dajani.

Education is a big topic for the United Nations at the moment. Across the world, officials, scientists and higher-education experts are discussing how to update the organization's Millennium Development Goals. One idea is that the proposed replacement, the Sustainable Development Goals, should widen the focus from school-age learning to the quality of higher education. Many of *Nature's* readers have experienced university teaching; many deliver it. All should have an opinion on it.

A high-level group set up by the European Commission to report on the quality of teaching in the region's higher-education institutions concluded that it is "an embarrassing disappointment". The report, published in June, added: "Serious commitment to best practice in the delivery of this core teaching mission is not universal, is sporadic at best and frequently reliant on the enlightened commitment of a few individuals." Similar concerns have been expressed about higher education in the United States.

The problem becomes more acute when set against the backdrop of continued economic uncertainty. As Chinese scientist Qiang Wang pointed out on this page earlier this year (see *Nature* 499, 381; 2013), there is a worrying disconnect between what is taught in schools and universities, and the skills those students need in the real world. He was talking about China, but students everywhere face the same conundrum. We need to educate our youth to be entrepreneurs, so that they can create their own jobs despite the economic uncertainty.

Students in the Arab world also face political uncertainty, as demonstrated by the events of the 'Arab Spring' and the continuing tensions. Despite some progress in education — in school enrolment, for example — the need for innovative teaching strategies is even more urgent here. Education reform in these countries tends to focus on the construction of new buildings, facilities and curricula. Knowledge, information and theories are presented as indisputable facts, and this creates students who struggle with the idea of uncertainty and who do not develop the analytical and problem-solving skills they need to prosper.

We need to shape and develop our own education systems. Simply reproducing Western models of education runs the risk of, among other issues, ignoring the configurations of politics, religion and gender unique to our region. Indeed, once the Arab world expands its innovative educational programmes, a healthy synergy between East and West can develop.

In my teaching of cell biology to university students in Jordan, I have introduced some innovations aimed at making the students think for themselves. As we all know, the media often

inaccurately report science. I ask students to identify an item on the radio, television or in the newspapers, and to check whether it is true. Then they write to the media organization to outline their findings and add a note about the impact of misleading information on patients and the general public, and the importance of making the source of the story clear.

This is an example of what educators call 'service-learning'. The students learn through their own research, while simultaneously serving the community, in this case the media — which in theory could alter the way science is reported in the future. Service-learning lets students learn more than the facts; they discover the relevance of that knowledge to real life and how it affects a community. They see their role in building that community and acquire a sense of responsibility. When they graduate,

they have more confidence to try to drive change, even in a world of unemployment or instability.

Many university students are not interested in some courses they take, often because they are obligatory. I see it with some students on my molecular-biology course. To pique their interest, ideally I would like to give them a relevant novel to read, say *Darwin's Radio* by Greg Bear. As well as covering the basic concepts of molecular biology, this book discusses the ethics of its application to real-world situations. Classroom discussion would be enlivened by discussion of the characters and themes, and the students would develop different points of reference for looking at a particular issue. Drama can also be used to teach biological mechanisms. Involving students personally in a three-dimensional world makes them think of

the mechanism from the perspective of the molecule. They can then better understand the limitations, challenges, potential and beauty of cells.

Our societies do not need students who are merely textbook educated; we need students who can engage positively with society. Too often, higher education focuses on the former without paying attention to the latter. We are all potential entrepreneurs in the sense that we can easily identify problems. The bigger challenge, and where conventional education fails, is to enable us to overcome doubts and inhibitions and take action. The goal of higher education should be for students to learn to apply the knowledge and skills they acquire to the realm of everyday life.

As the poet William Butler Yeats said: "Education is not the filling of a pail, but the lighting of a fire." Our objective in education must be to light a fire in the heart of every individual. ■

Rana Dajani is assistant professor of molecular biology at the Hashemite University in Zarqa, Jordan, and former Fulbright visiting professor at Yale University.
e-mail: rdajani@hu.edu.jo

OUR OBJECTIVE IN
EDUCATION
MUST BE TO
LIGHT A FIRE
IN THE HEART
OF EVERY
INDIVIDUAL.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/itmhji

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

PHYSIOLOGY

Skin cells have daily rhythms

Stem cells from human skin keep to a 24-hour schedule that might protect them from sun damage.

Salvador Aznar Benitah, then of the Centre for Genomic Regulation in Barcelona, Spain, and his colleagues analysed cultures of genetically identical stem cells from human skin at set times. They found that genes related to the 'body clock' are expressed in distinct waves over a 24-hour cycle.

Each wave is associated with peaks in expression for other genes: those that protect against DNA-damaging sunlight are most active during the day, as are those involved in DNA replication and cell growth. Genes that push stem cells to become specialized are most active in the evening and night. Disruptions to the internal clock could lead to premature ageing, the researchers suggest. *Cell Stem Cell* <http://doi.org/pbb> (2013)

CONSERVATION

Florida panthers keep their heads

Endangered Florida panthers have maintained their distinctive faces despite cross-breeding.

Human activity in the twentieth century drove this subspecies of *Puma concolor* (pictured) towards extinction and confined it to the southern tip of the Florida peninsula.



To combat severe inbreeding, eight Texas pumas were temporarily introduced to mate with this population.

David Reed at the Florida Museum of Natural History in Gainesville and his colleagues analysed the skulls of 20 male and 20 female panthers to see whether this cross-breeding had affected the animals' distinctive facial features. They found that identifying characteristics such as a highly arched 'Roman nose' have not been significantly altered. Those panthers born from crosses with their Texas cousins were similar to 'pure' Florida animals. *J. Mammal.* 94, 1037–1047 (2013)



ECOLOGY

Counting trees in the Amazon

The Amazon rainforest is renowned for its biodiversity, but just 227 'hyperdominant' species account for half of all trees across the 6-million-square-kilometre basin.

Hans ter Steege of the Naturalis Biodiversity Center in Leiden, the Netherlands, and his colleagues analysed data from 1,170 plots scattered across the forest and then extrapolated their data to the entire basin. They calculated

that the region contains around 390 billion trees with trunks of 10 centimetres or more in diameter, and some 16,000 species.

The authors suggest that the extreme dominance of a few species could simplify efforts to understand the large-scale ecology of the basin, but might complicate efforts to identify rare species that are at risk of extinction. *Science* <http://doi.org/pb2> (2013)

BIOCHEMISTRY

Icy origins for RNA copying?

For the first time, experiments in evolution have produced an RNA molecule that can build other RNA molecules that are longer than itself.

Many theories of the origin of life rely on RNA self-replication, but researchers have struggled to make RNA 'enzymes' that can stitch together other RNAs of a similar size. Reasoning that freezing temperatures would stabilize RNA synthesis, Philipp Holliger and his colleagues at the Medical Research Council Laboratory

of Molecular Biology in Cambridge, UK, ran *in vitro* evolution experiments in ice, producing RNA enzymes that can synthesize RNA at temperatures as low as -19°C in tiny pockets between ice crystals.

By combining cold-generated mutations with those from previous work, the researchers created the most-efficient RNA enzyme so far: a 202-nucleotide molecule that can copy templates as long as 206 nucleotides. Ice could have aided the emergence of self-replication in the prebiotic chemical world, the authors say. *Nature Chem.* <http://doi.org/pcs> (2013)

DANIEL SABATIER/IRD/UMR/AMAP

MARK CONLIN/ALAMY

SEVEN DAYS

The news in brief

EVENTS

US shutdown ends

People in the United States breathed a collective sigh of relief on 17 October when a last-minute budget deal between lawmakers reopened the government, which had shut down on 1 October. A stopgap measure will fund government operations until 15 January. Science agencies are now scrambling to restart research programmes. See page 419 and go.nature.com/x9swmx for more.

POLICY

Energy storage

The California Public Utilities Commission has approved the first energy-storage plan in the United States. Adopted on 17 October, the scheme promotes the use of renewable sources such as wind turbines and solar panels, which produce energy intermittently. Under the regulation, three major utility companies must buy a combined 200 megawatts of energy-storage capacity by 1 March 2014, and a total of 1,325 megawatts of storage by 2020.

Science for UN

Twenty-six scientists from around the world have been appointed to a newly created Scientific Advisory Board for the United Nations. The panel is charged with providing science-based advice on environmental, developmental and socio-ethical issues. See go.nature.com/4ts2qb for more.

Spanish bailout

Spain's science system received a much-needed cash infusion on 18 October, when the government approved a €70-million (US\$96-million) package to save the Spanish National Research Council

(CSIC) from bankruptcy. In June, the council received an extra €25 million in government support, but in July, CSIC president Emilio Lora-Tamayo said that a further €75 million would be needed by the end of the year. As Spain's largest scientific organization, the CSIC maintains more than 100 institutes and supports about 6,000 scientists. See go.nature.com/gesitc for more.

RESEARCH

Smoking gun

The *British Medical Journal* (BMJ) announced on 15 October that it will no longer publish studies funded by the tobacco industry. BMJ editors had previously defended the

inclusion of such studies but reversed course last week, citing the industry's wilful misuse of research to cast doubt on the health risks of smoking. The policy applies to the *BMJ* and its sister journals *Thorax*, *Heart* and *BMJ Open*. The American Thoracic Society already refuses tobacco-industry-funded studies, as do some journals published by the Public Library of Science.

Gravity mission

Europe's gravity-hunting space mission is over. Having run out of xenon fuel, the Gravity Field and Steady-state Ocean Circulation Explorer (GOCE) will re-enter Earth's atmosphere within weeks, the European Space Agency said on 21 October. GOCE has

produced the most accurate gravity maps yet — in part because its final measurements were taken from an unusually low orbital altitude of 224 kilometres. Since its launch in 2009, the mission has created many maps, including records of ocean circulation and the planetary gravitational reference known as the geoid (see *Nature* 458, 133; 2009).

Reproducibility test

An initiative to replicate important research results has been awarded US\$1.3 million to verify 50 high-profile cancer studies from the past three years. The Reproducibility Initiative, co-founded by Elizabeth Iorns (see *Nature* 500, 14–16; 2013), will repeat studies including 27 published



ALEXANDER FIRSOV/AP

Russian lake yields massive meteorite

On 16 October, Russian scientists recovered a large chunk of the 9,000-tonne meteorite that exploded over the Ural region in February, injuring more than 1,000 people (see *Nature* <http://doi.org/pck>; 2013). Weighing around 600 kilograms, the blackish rock (pictured) was winched out of Lake Chebarkul. Reconstructions of the meteorite's trajectory

— and an ominous hole in the frozen surface of Lake Chebarkul on the morning after the impact — led scientists to suspect that the main fragment had landed there (see *Nature* 495, 16–17; 2013). “This is without doubt the largest fragment yet of the Chelyabinsk meteorite,” says researcher Viktor Grokhovsky of the Ural Federal University in Ekaterinburg.

CIRM

in *Nature*. The grant, from the Laura and John Arnold Foundation in Houston, Texas, was announced on 16 October. See go.nature.com/bqxm5q for more.

Ethics update

The World Medical Association has revised the Declaration of Helsinki, an influential guide to ethical conduct in research on human subjects. The international association of physicians, based in Ferney-Voltaire, France, approved the updated version on 19 October. The revision includes provisions for compensating people who are harmed in the course of research, strengthens protection for vulnerable populations and renews the association's call for the sharing of research results.

PEOPLE

Stem-cell leader

Alan Trounson (pictured) will step down as president of the California Institute for Regenerative Medicine (CIRM) in San Francisco, the organization announced on 16 October. Trounson, a former stem-cell scientist at Monash University in Melbourne, Australia, joined the CIRM in 2007. His replacement will be charged with navigating the publicly funded agency, which was established in 2004 with a



US\$3-billion allocation from the state, through an uncertain financial future (see *Nature* **482**, 15; 2012).

Data faked

Nitin Aggarwal, a cardiac scientist formerly of the University of Wisconsin–Madison, has agreed to have his research supervised for the next three years, and to be excluded during that time from peer-review committees for US agencies such as the National Institutes of Health (NIH). The US Office of Research Integrity reported on 17 October that Aggarwal had falsified or fabricated data in his graduate thesis, two journal articles and grant applications to the American Heart Association and the NIH.

BUSINESS

End sequence

Roche, a health-care company based in Basel, Switzerland, confirmed last week that it

will discontinue its 454 Life Sciences sequencing platform in 2016. The platform has struggled to compete with cheaper, more accurate alternatives since being acquired by Roche in 2007. The company said last week that about 100 employees will be laid off when it closes its facility in Branford, Connecticut. Roche ended internal research-and-development efforts on third-generation sequencing technologies in April, and in September signed a US\$75-million deal to develop diagnostics applications with Pacific Biosciences, based in Menlo Park, California.

FUNDING

Ocean monitoring

The US National Science Foundation announced a US\$16-million award on 18 October to launch an ocean-observing array in the North Atlantic. Deep-water currents in that region are part of a global system that is thought to affect weather and climate (see *Nature* **497**, 167–168; 2013). Disbursed over five years, the money will fund the Overturning in the Subpolar North Atlantic Program — a multinational effort to monitor ocean temperature, salinity and the strength of currents along a line that runs from

COMING UP

27–30 OCTOBER

Extreme rain and floods that hit Colorado in September are discussed at the Geological Society of America annual meeting in Denver. go.nature.com/7quicy

30 OCT–1 NOV

Topics range from health to agriculture at the 8th International Conference on Genomics in Shenzhen, China. The programme also highlights big-data management and open data platforms. go.nature.com/qgpvtl

Newfoundland in Canada to Scotland, passing Greenland's southern tip.

African genetics

Genomics research in Africa received a boost on 18 October when the US National Institutes of Health (NIH) in Bethesda, Maryland, announced the award of ten grants totalling US\$17 million from the Human Heredity and Health in Africa (H3Africa) programme. The four-year grants will fund research on the role of genetics in disorders such as tuberculosis and African sleeping sickness, and will support two science centres in Nigeria. Backed by the NIH and the UK Wellcome Trust, H3Africa has awarded \$74 million for research since its inception in 2010.

CORRECTION

The story 'Nobel laureate dies' (*Nature* **501**, 467; 2013) should have said that neural signals generated from light, rather than light itself, are transmitted from the eye to the visual cortex.

NATURE.COM

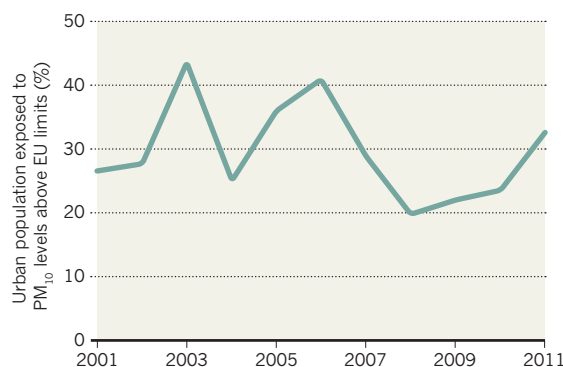
For daily news updates see: www.nature.com/news

TREND WATCH

The World Health Organization's cancer agency has classified outdoor air pollution as a human carcinogen. On 17 October, the International Agency for Research on Cancer cited studies linking dirty air to lung cancer and an increased risk of bladder cancer. The agency also labelled the particulate matter found in outdoor air pollution as a cause of cancer. Urban exposure to particulates in Europe was highlighted in a separate report last week by the European Environment Agency (see chart).

HOLD YOUR BREATH

Between 2001 and 2011, about one-third of Europe's city dwellers were exposed to hazardous levels of particulate matter in the air.



* PM₁₀, particulate matter smaller than 10 micrometres in diameter. EU limits: 50 micrograms PM₁₀ per cubic metre, not to be exceeded on more than 35 days a year.

SOURCE: EEA/ESTAT

NEWS IN FOCUS



POLITICS Lost data, delays and despondency in the wake of US shutdown **p.419**

PHYSICS 'Bated breath' as first results from dark-matter study are due **p.421**

ATMOSPHERE Giant fake cloud of volcano ash to test flight-safety system **p.422**

PALAEONTOLOGY The big, fierce enduring debates about *T. rex* **p.424**

JOE RAEDLE/NEWSMAKERS/GETTY



The lethal-injection chamber in Huntsville, Texas.

ANAESTHETICS

Death row incurs drug penalty

Bid to use common anaesthetic for executions threatens to cut off supply to US hospitals.

BY CHRIS WOOLSTON

Allen Nicklasson has had a temporary reprieve. Scheduled to be executed by lethal injection in Missouri on 23 October, the convicted killer was given a stay of execution by the state's governor, Jay Nixon, on 11 October — but not because his guilt was in doubt. Nicklasson will live a while longer because one of the drugs that was supposed to be used in his execution — a widely used anaesthetic called propofol — is at the centre of an international controversy that threatens

millions of US patients, and affects the way that US states execute inmates.

Shortages of anaesthetic drugs usually used in lethal injection, the most common method of execution, are forcing states to find alternative sedatives. Propofol, used up to 50 million times a year in US surgical procedures, has never been used in an execution. If the execution had gone ahead, US hospitals could have lost access to the drug because 90% of the US supply is made and exported by a German company subject to European Union (EU) regulations that restrict the export of medicines and devices

that could be used for capital punishment or torture. Fearing a ban on propofol sales to the United States, in 2012 the drug's manufacturer, Fresenius Kabi in Bad Homburg, ordered its US distributors not to provide the drug to prisons.

This is not the first time that the EU's anti-death-penalty stance has affected the US supply of anaesthetics. Since 2011, a popular sedative called sodium thiopental has been unavailable in the United States. The manufacturer, US company Hospira, abandoned plans to produce the drug at its plant in Italy after regulators in the country required that the thiopental never be used in executions. The drug, which is difficult and costly to make, was already in short supply because of manufacturing problems.

"There has been a collision of the politics of capital punishment in the United States and Europe, forcing us to hopscotch around looking for suitable methods for anaesthesia," says Jerry Cohen, a former president of the American Society of Anesthesiology.

"The European Union is serious," says David Lubarsky, head of the anaesthesiology department at the University of Miami Miller School of Medicine in Florida. "They've already shown that with thiopental. If we go down this road with propofol, a lot of good people who need anaesthesia are going to be harmed."

The loss of thiopental from the anaesthesia arsenal was a relatively minor inconvenience, says Cohen, because propofol provided an alternative. But if propofol is used for executions in Missouri or any other state, it could disappear too, leaving hospitals in a serious bind. "Propofol has a lot of uses for which there are no substitutes," says Cohen. It is the preferred way to sedate people who have breathing tubes because it acts quickly and does not cause vomiting. Federal regulations make propofol difficult to manufacture in the United States.

The 35 US states with prisoners on death row were already scrambling to find effective drugs for lethal injection, which was used for 43 executions last year. The procedure previously relied on a course of three injections: thiopental to sedate the prisoner, muscle relaxant pancuronium bromide to induce paralysis, and potassium chloride to stop the heart. As supplies of thiopental ran low in 2009 and 2010, many states started stockpiling pentobarbital, another sedative. But in 2011, Lundbeck, a drug company in Copenhagen and sole US supplier of pentobarbital, banned it from use in executions because of Danish and EU human-rights ►

► laws. Texas's supply of pentobarbital expired in September, but the state obtained more from unregulated compounding pharmacies, which tailor-make drugs. Pentobarbital is not "especially" useful as a surgical anaesthetic, says Lubarsky, so its shortage has little impact on patient care.

On 15 October, after running out of pentobarbital, Florida executed William Happ using midazolam as the sedative. But midazolam, which is similar to diazepam (Valium), had never been used in an execution, and, according to media reports, Happ was still blinking and moving his head minutes after the injection.

Nobody knows whether midazolam is appropriate for lethal injections, says Lubarsky. "We've turned this into a circus of experimenting on prisoners," he says. "The state is playing doctor without any regard for efficacy. It changes protocols willy-nilly." The drug is not a good anaesthetic, he says, and it may not shield prisoners from the pain of the final injection.

Although midazolam has now entered the realm of capital punishment, it is unlikely that surgical supplies will be affected. Hospira is one of many companies that makes midazolam and has no plans to stop, says Dan Rosenberg, a company spokesman. Rosenberg would not

say where Hospira makes midazolam, but he says that European regulations "aren't an issue".

Meanwhile, Missouri has suspended another execution, scheduled for 20 November, while it tries to find an alternative to propofol. Lubarsky notes that although a single, large dose of propofol could work as a method of execution, its use in US prisons would be problematic because it could be complex to administer and physicians are generally not willing to participate in the process (see *Nature* **441**, 8–9; 2006). "Putting together a foolproof protocol that could be carried out by prison guards with high-school educations is another matter entirely," he says. ■

PUBLISHING

Brazil fêtes open-access site

South American SciELO project weighs up future after 15 years of free publishing.

BY RICHARD VAN NOORDEN

Researchers and publishers are gathering this week in São Paulo, Brazil, to celebrate a quietly subversive open-access publishing project. The occasion: the 15th anniversary of SciELO (Scientific Electronic Library Online), a subsidized collection of mainly Latin American journals that now puts out more than 40,000 free-to-read articles each year — and which aims to put developing countries firmly on the scientific map.

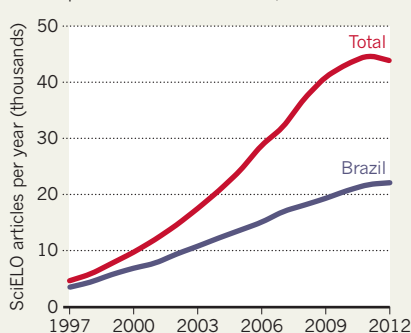
Although little noticed by European and North American scientists, SciELO is "one of the more exciting projects not only from emergent countries, but also in the whole world", argues Jean-Claude Guéron, an open-access supporter who studies comparative literature at the University of Montreal in Canada.

In contrast to fee-charging open-access journals, journals on the SciELO platform charge authors little or nothing to publish because state and government funders provide infrastructure and software. That backing has helped to make Brazilian research the most open in the world — in 2011, 43% of Brazilian science articles were free to read on publication, compared with, for example, 6% of US articles.

But on its 15th birthday, SciELO's future is in flux. Broader recognition of the venture might inspire similar 'public-good' networks in other emerging science regions. Or the project might dwindle in influence as commercial open-access publishers muscle in. "The direction that SciELO goes in will have a big effect on scholarly communications in Latin America," says Juan Pablo Alperin, a doctoral student at Stanford University in California who develops software at the Public Knowledge Project, a research initiative

FREE AND EASY

SciELO has expanded rapidly. For comparison, the global number of immediately available open-access articles published in 2011 was 340,000.



looking at open-access scholarly publishing.

The roots of SciELO go back to 1993, when Rogério Meneghini, now SciELO's scientific director but then at the São Paulo Research Foundation (FAPESP), saw that "a great deal of [Brazil's] scientific conversation was not noticed in global science". In an effort to raise the visibility of Brazilian research, FAPESP started funding SciELO as a one-year pilot project in 1997, with journals that met basic editorial standards being placed in the collection. Ten other countries, including Mexico, Spain and South Africa, subsequently joined. And it has inspired other free Ibero-American publishing platforms, such as the 11-year-old Redalyc.org.

Much of the project is funded by a US\$3-million annual grant from FAPESP and from Brazil's National Council for Scientific and Technological Development, says SciELO director Abel Packer. Separately, some journals offer extra services, such as English translation. And each country supports its own journal

operations (South Africa, for example, has chipped in with \$450,000; Chile, with \$345,000).

SciELO's admirers say that the system builds publishing expertise and helps researchers to publish open science on regional subjects — such as health issues and farming techniques — that might be rejected by international journals. However, citations are low and journal quality variable. Many Brazilian researchers choose instead to publish in international journals, notes Margareth Capurro, a biologist at the University of São Paulo. This is partly because funding agencies prefer higher-impact publications, she adds.

"If 'influence' were measured by other ways, such as usage, we may see a different picture," says Leslie Chan, who studies open access at the University of Toronto in Canada. SciELO Brazil gets 1.5 million downloads per day, and this year, a SciELO citation database will be added to the Thomson Reuters Web of Knowledge, further raising visibility.

Packer and Meneghini hope to persuade other emergent science nations to join: India has been approached. They say that, for the Brazilian journals, the greatest challenges are to raise journal quality and international recognition. This might involve professionalizing editorial boards and paying salaries. But that could mean higher costs, says Meneghini.

As SciELO grows (see 'Free and easy'), its biggest journals are in danger of being bought by profit-seeking publishers, warns Guéron. That would be a shame, Alperin says, adding that a free-to-publish system helps to sidestep problematic aspects of open-access publishing, such as when fee-charging journals accept as many papers as possible without providing adequate peer review. "I'd love to see more of the world copy the Latin American model," he says. ■

SciELO

POLITICS

Pain of US shutdown lingers

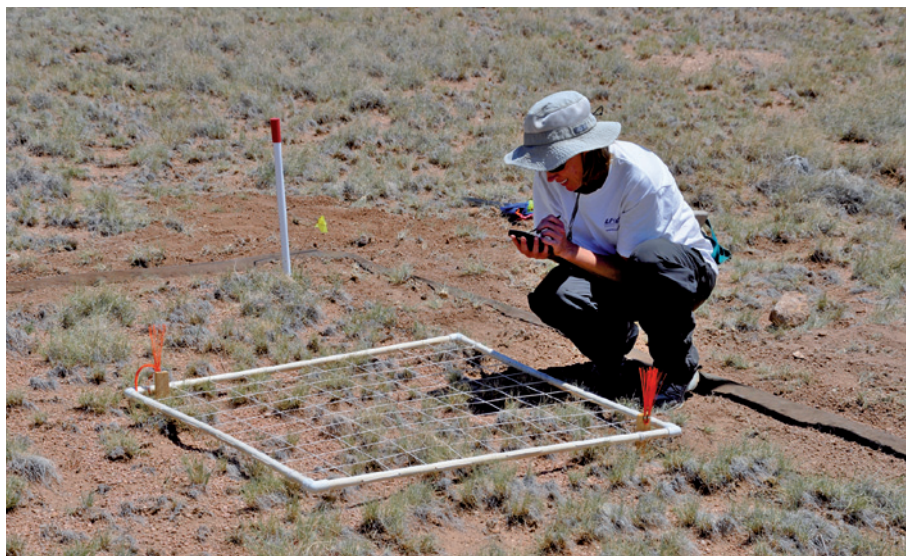
Researchers fear that continuing budget fights will further harm government-funded science.

BY LAUREN MORELLO, HEIDI LEDFORD,
HELEN SHEN, JEFF TOLLEFSON,
ALEXANDRA WITZE & SARAH ZHANG

Ecologist Stacy Kim should be preparing to leave for Antarctica, where she was due to begin a study of marine life in the Ross Sea next month. Instead, she is trying to work out how to keep her lab running after her polar plans were cancelled by the US National Science Foundation (NSF), which is struggling to salvage a field season shortened by the 16-day US government shutdown that ended on 17 October.

It is not only the loss of a potential year's worth of data that pains Kim, a researcher at the Moss Landing Marine Laboratories in California, who hoped to use a remotely operated underwater vehicle to monitor everything from Ross Sea phytoplankton to killer whales. The NSF's decision also jeopardizes the flow of grant money to her lab, and she may be forced to lay off technicians. "Most of the people in my lab group have sublet their places for the three months we were supposed to be in Antarctica," she says. "Now I have homeless people," who, she adds, may have to go on unemployment benefits.

But the worst may not be over for US researchers, who face the possibility of another government shutdown in mid-January, when the deeply divided US Congress must agree on a new plan to fund government operations. "I don't think we've learned anything" from the last shutdown, says Matt Hourihan, who directs the research and development



SCOTT COLLINS

The shutdown disrupted long-term monitoring projects such as a survey of plant life in New Mexico.

budget and policy programme at the American Association for the Advancement of Science (AAAS) in Washington DC. "In many ways, the big fiscal challenges are still there."

These include not only the threat of another shutdown — which would again bring US science agencies' research and grant-making to a halt — but also the scheme known as sequestration. This took a 5.1% bite out of most US government programmes this year and is poised to claim still more in January. Overall, federal spending on research and development has dropped by an astounding 16.3% since 2010, according to a recent AAAS analysis, and Congress often chooses to fund the government with temporary spending plans (see 'Passing the buck').

"If I were a young person today, I'd have to wonder if I'd want to go into science in the United States anymore, because the uncertainty has become extraordinary," says Michael Lubell, director of public affairs for the American Physical Society in Washington DC.

At Moss Landing, Kim is struggling to help her students cope with a year's delay to their research plans — an especially difficult task for those pursuing two-year master's degrees, and for an incoming doctoral student. But it is not only young scientists who have lost valuable time and data.

For more than two weeks, Scott Collins, a biologist at the University of New Mexico in Albuquerque, could not access his research site in the Sevilleta National Wildlife Refuge, which closed during the shutdown. That short absence

will make it harder to understand how this year's unusually rainy summer affected plant life in arid New Mexico. "We're funded with federal tax dollars to do this research," Collins says. "It matters to us that we do the job well."

And, across the country at a US Department of Agriculture facility in Newark, Delaware, entomologist David Jennings returned to work last week to find that many of his colonies of emerald ash borer larvae had perished. The small crew of 'essential' personnel left to run the lab during the shutdown could not maintain the temperature and feeding schedule that the picky larvae require. Jennings estimates that it will take close to a year to recoup what the lab has lost, delaying research on how to protect US forests from the tree-munching beetle.

Also in jeopardy are some major infrastructure projects currently in development. For example, the shutdown has postponed the final design review for the Large Synoptic Survey Telescope, a ground-breaking project that would allow astronomers to map the southern sky once every three days. The delay could prevent construction from starting next year as planned.

Pieter Tans, who heads the Carbon Cycle Greenhouse Gases group at the National Oceanic and Atmospheric Administration lab in Boulder, Colorado, says that the political manoeuvring that caused the shutdown will also have long-term effects on morale at US research agencies. "It implicitly sent the message to the American people that they don't need all of these government scientists," he says. ■ [SEE COMMENT P.431](#)

PASSING THE BUCK

The US Congress has relied heavily on stopgap spending measures to keep the government running.





The fishing industry has been preparing for a key European parliamentary vote on subsidies.

FISHERIES POLICY

Europe debates fisheries funding

Campaigners want subsidies to be focused on conservation.

BY DANIEL CRESSEY

As *Nature* went to press, the European Parliament was voting on how billions of euros in subsidies should be allocated to the fishing industry. In past years, the main focus has been on ‘capacity building’ — the strengthening and support of fishing fleets. But now, after years of worries about overfishing and damage to the marine environment, calls are growing among scientists for more spending on sustainability and conservation.

The battle lines have been drawn. Some of the roughly €6.4 billion (US\$8.7 billion) in subsidies earmarked to support fishing between 2014 and 2020 — known as the European Maritime and Fisheries Fund — is slated to go to conservation and data collection. But, as in the past, much of the money could be spent on modernizing vessels, cutting fuel costs and even on the construction of fishing boats.

These measures would please fishermen but outrage conservation groups and some scientists, who fear that a vote by Members of the European Parliament (MEPs) to subsidize an increase in fishing capacity could undo work to put fishing on a more scientific footing. Europe has long been criticized for ignoring advice on safe levels of fishing, but this year the European

Union (EU) took a big step forward when it agreed a package of legislation to put science at the centre of all decisions on setting catch quotas (see *Nature* 498, 17–18; 2013). Voting for capacity-enhancing subsidies could undermine that achievement, campaigners argue.

Researchers also point out that Europe catches more fish than is sustainable in many areas. By the European Commission’s own estimates,

four-fifths of Mediterranean fish stocks and almost half of Atlantic stocks are overfished, leaving populations of species such as cod and mackerel in a bad way. Subsidizing fleets to boost catches could be devastating to ecosystems that are already under pressure, critics say.

Ahead of the vote, a campaign by researchers has challenged MEPs to amend the funding legislation so that subsidies instead go to better management and research, such as assessments of how many fish are in the seas, the setting up of marine reserves and basic oceanographic studies. More than 180 researchers have signed a letter urging MEPs to support this measure.

Rashid Sumaila, director of the Fisheries Economics Research Unit at the University of British Columbia in Vancouver, Canada, was one of the organizers of the letter and the lead author of a report submitted to Parliament last week. In it, he and his colleagues estimate that about \$35 billion is spent on subsidies globally each year, with capacity-enhancing subsidies making up more than \$20 billion of that (see go.nature.com/yxpfe2 and ‘Net spend’). Sumaila and his colleagues want an end to payments that increase the ability of fishing fleets to catch fish, including those that cut fuel costs and fund the modernization of boats.

Sumaila admits that these recommendations will not “go down well” with politicians and fishermen. But, he says, “if sustainable fisheries are your goal, you need to cut the subsidies”.

The vote is being watched closely. Some nations — notably New Zealand — have made moves to phase out damaging subsidies, but a similar global agreement has proved harder to achieve. Many countries, such as France and Spain, are wedded to subsidies, which they believe support a crucial food sector.

And the dispute could have repercussions for global trade. Fisheries subsidies are being discussed at the World Trade Organization, but those talks are deadlocked. The EU has repeatedly said that it supports the elimination of subsidies that contribute to overcapacity. A vote in the other direction now could make it harder to get global agreement. “On an international level, people are always watching the EU,” says Markus Knigge, a policy expert at the Brussels-based European Marine Programme run by the Pew Charitable Trusts.

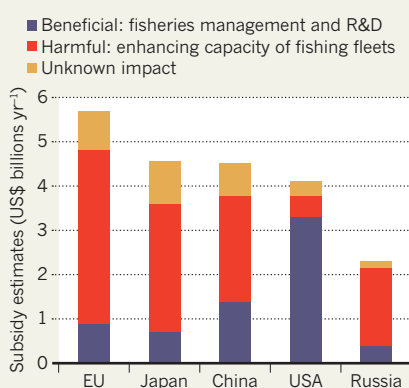
Once it has voted, the European Parliament will enter into negotiations with the European Council — made up of representatives of the EU’s 28 member states. A final agreement on the subsidy package is expected early next year.

Ray Hilborn, a fisheries researcher at the University of Washington in Seattle, argues that Europe already has a well-developed management system for its fisheries. “If they would just keep the politicians out of quota setting, they would do pretty well,” he says.

And, he adds, a properly managed fishery should not need subsidies: “If fisheries are well managed, they are very profitable and they should have to fend for themselves.” ■

NET SPEND

Researchers say that fisheries subsidizers allocate more to potentially harmful subsidies such as fuel than to ‘beneficial’ activities such as conservation.



PHYSICS

Final word is near on dark-matter signal

An influential US experiment prepares to release its first results.

BY EUGENIE SAMUEL REICH

Viewed end on, the arrays of photomultiplier tubes on the Large Underground Xenon (LUX) experiment look like beds of flowers. The hope is that they will capture sparks of light emitted when particles of dark matter collide with liquid xenon. With 122 detector tubes, LUX is much more sensitive than its closest rival in the competitive field of dark-matter searches — and in just days, physicists the world over will know whether that advantage has yielded definitive results.

The project, based at the Sanford Underground Research Facility in Lead, South Dakota, will release its first findings on 30 October. They are likely to reveal whether tentative dark-matter signals seen by other experiments are real, and will also inform ongoing discussions about how much more time and money should be spent on the hunt for dark matter. “The potential is there, and all the community is waiting with bated breath to see what they observe,” says Juan Collar, a physicist who leads a rival experiment at the University of Chicago in Illinois.

Elena Aprile, a physicist at Columbia University in New York city who leads another competitor, XENON100, based at Gran Sasso National Laboratory near L'Aquila, Italy, is betting that LUX has not seen dark matter. “A null result is all that can be expected at this stage,” she says. A LUX spokesman, physicist Daniel McKinsey of Yale University in New Haven, Connecticut, says simply: “We have a detector that is working very, very well.”

LUX came online this year amid fierce debate. Scientists know from astronomical observations that five-sixths of the matter in the Universe is dark — making itself known mostly through its gravitational tug on bright matter — but attempts to detect it directly, on its presumed passage through Earth, have been fraught with controversy.

The DAMA/LIBRA experiment (Dark Matter Large Sodium Iodide Bulk for Rare Processes) at Gran Sasso reported a statistically significant signal more than 10 years ago, but physicists have not independently confirmed the result. In 2010, the Coherent Germanium Neutrino Technology experiment in Soudan, Minnesota, and the Cryogenic Dark Matter Search



LUX could resolve a decade-long physics debate.

at the University of California, Berkeley, each reported tantalizing, but not statistically convincing, glimpses of potential dark matter; a year later, XENON100 saw no sign of the stuff. That prompted heated discussion over whether the experiment was sensitive to the lighter dark-matter particles that might have been glimpsed by the other two experiments.

Enter LUX, which will deliver its first results just as the US Department of Energy decides which of several dark-matter experiments should be given money to expand. LUX wants to install a larger, 7-tonne detector, in a proposed US\$30-million project called LUX-Zeplin. McKinsey argues that such experiments should be scaled up until they hit a physical limit — when the background noise from other weakly interacting particles becomes overwhelming. “That’s a natural break point,” agrees Jonathan Feng, a theoretical physicist at the University of California, Irvine.

One candidate for dark matter is the neutralino, a particle predicted by some supersymmetric theories of particle physics, in which particles are paired with heavier counterparts. If, as Feng expects, LUX sets a detection threshold around three times more stringent than that of XENON100, it will rule out some types of neutralino. “There’s an unbelievable amount of effort focused on the neutralino, so this upcoming announcement is quite important,” he says. ■

LUX/DARKMATTER



The 2010 eruption of Eyjafjallajökull grounded aircraft with engines that are vulnerable to volcanic ash.

ATMOSPHERIC SCIENCE

Volcanic-ash sensor to take flight

Researchers will fly jet towards giant artificial particle cloud to test safety device.

BY ALEXANDRA WITZE

On 28 October, if all is calm and clear off the west coast of France, Fred Prata will help to simulate a near-disaster. Prata, an atmospheric scientist at Nicarnica Aviation in Kjeller, Norway, has planned the biggest field test yet for a device intended to help aeroplanes to survive close encounters with volcanic ash, which can melt in the high temperatures of jet engines and form a glassy coating that chokes airflow.

Instead of an actual erupting volcano, Prata and his team have a tonne of ash, flown in

from the Icelandic volcano Eyjafjallajökull. And instead of Europe's aviation industry, they have a jet plane that will fly towards an artificial cloud of that ash. The goal is to test an infrared camera that alerts pilots to volcanic particles in their path.

Prata has been trying to get his sensor onto jets since he first developed it more than 20 years ago (A. J. Prata *et al. Nature* **354**, 25; 1991). He had only moderate success until the 2010 eruption of Eyjafjallajökull sent ash into European airspace and grounded flights for nearly a week, prompting the airline carrier easyJet and the manufacturer Airbus to invest in

his efforts at Nicarnica, an offshoot of the Norwegian Institute for Air Research. This month's test could be a major step towards getting the sensor onto commercial jets worldwide.

The work highlights how much scientists have learned about volcanic ash since Eyjafjallajökull brought much of Europe to its knees. The eruption "brought different disciplines together in ways that weren't integrated before", says Sue Loughlin, head of volcanology for the British Geological Survey in Edinburgh, UK. "That's been a really great thing." What these researchers learned has led European regulators to devise new guidelines on how much ash is acceptable for planes to fly through. And scientists have improved their understanding of how the spread of ash over long distances is affected by factors such as weather patterns.

Prata's sensor, the Airborne Volcanic Object Imaging Detector (AVOID), uses infrared cameras to detect the silicate particles in volcanic ash. In 2011, it flew in successful low-elevation tests at Italy's erupting Etna and Stromboli volcanoes. The upcoming experiment will involve the largest artificial ash cloud ever made, and will probably be over the Bay of Biscay, in airspace controlled by the French military. (There is a backup site on France's Mediterranean coast in case of bad weather.)

An Airbus A400M cargo plane will fly in a tight spiral, dispensing ash from 50 barrels as it climbs from 3,000 metres to almost 4,000 metres (see 'Silver lining'). A second plane, an Airbus A340 commercial airliner carrying the AVOID sensor, will fly near the cloud at various heights, taking measurements. A four-seater propeller plane from the Düsseldorf University of Applied Sciences in Germany will measure optical properties from inside the cloud. Without a jet engine, this plane is not at risk of engine failure; it has previously flown in heavy ash plumes above active volcanoes, says Konradin Weber, leader of the Düsseldorf team.

At its densest, the artificial cloud is likely to contain no more than 1 milligram of ash per cubic metre, says Prata. That puts it at the low end of air contamination under European regulations adopted after Eyjafjallajökull. Anything below 0.2 milligrams is considered safe to fly in; between 0.2 and 2 milligrams, a pilot must be aware of ash hazards; between 2 and 4 milligrams, a pilot must conduct a special risk assessment to fly; and above 4, all flights are grounded.

It is not clear whether the artificial ash cloud

KJARTAN THORBJORNSSON/MORGUNBLADID/POLARIS/EVEVINE



MORE ONLINE

TOP STORY



Kepler finds two exoplanets orbiting at a steep angle to their star's equator go.nature.com/liu2bd

MORE NEWS

- HIV vaccine raised infection risk go.nature.com/a36gz4
- Spain saves research council from imminent bankruptcy go.nature.com/gesitc
- Fossil skull suggests that *Homo erectus* should subsume two other hominin species go.nature.com/h4gnpl

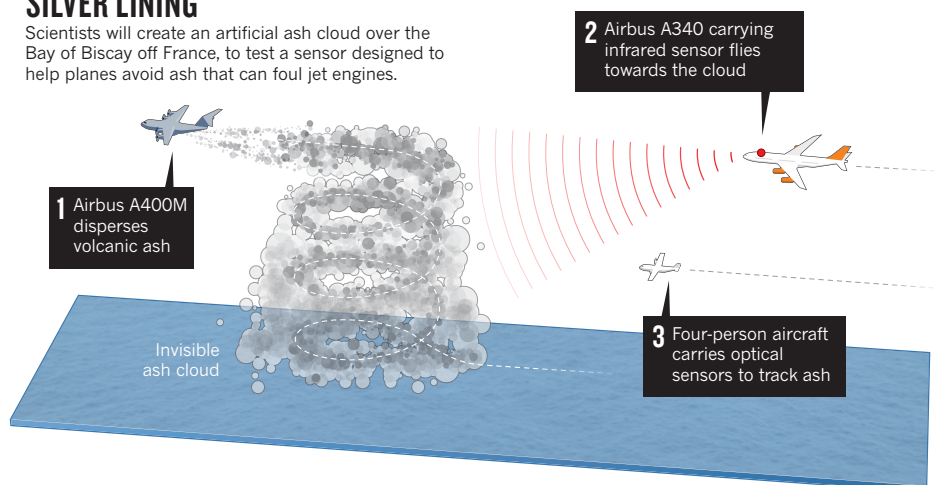
NATURE PODCAST



Reading minds by decoding brains; the truth about *T. rex*; and the week's top news in science nature.com/nature/podcast

SILVER LINING

Scientists will create an artificial ash cloud over the Bay of Biscay off France, to test a sensor designed to help planes avoid ash that can foul jet engines.



will be visible to the human eye, although scientists on a German research jet did spot Eyjafjallajökull ash in 2010, at concentrations below 0.2 milligrams of ash per cubic metre (U. Schumann *et al. Atmos. Chem. Phys.* **11**, 2245–2279; 2011). The artificial cloud is likely to dissipate in 6 to 12 hours, falling out harmlessly over the ocean, says Prata. The experiment will cost roughly €500,000 (US\$680,000) and, he says, “We have only one shot.”

The researchers will know just how much ash is released, and its precise geometry, so the

experiment will provide the best test yet for AVOID. But many hurdles remain before the system can be used commercially, including the need to integrate it into a working cockpit, and to scale up production. “It’s really not clear what we will do next,” says Prata. The decision rests mostly with Airbus, which would need to decide whether to develop the technology further. Prata hopes that AVOID could one day be used on planes flying in volcanically active regions from Indonesia to Chile or Alaska.

Back where it all began, a major

initiative called FUTUREVOLC is focusing on improving monitoring of Icelandic volcanoes. Led by the University of Iceland in Reykjavik and the Icelandic Meteorological Office, researchers are beefing up networks of equipment including seismic stations, cameras and gas detectors. “We’re working on all aspects, from magma generation inside the crust to how it progresses into eruption plumes and how this is dispersed,” says Freysteinn Sigmundsson, an earth scientist at the University of Iceland and co-coordinator of the project.

Even Prata is involved in FUTUREVOLC: he plans to deploy three of Nicarnica’s infrared cameras on the ground in Iceland. They will measure how fast and how high ash plumes rise — on their way to disrupting airspace somewhere. ■

CORRECTIONS

The News story ‘Study aims to put IPCC under a lens’ (*Nature* **502**, 281; 2013) said that Jean-Pascal van Ypersele was at the Catholic University of Leuven. He is at the Catholic University of Louvain in Louvain-la-Neuve. The Editorial ‘The maze of impact metrics’ (*Nature* **502**, 271; 2013) wrongly located the University of North Texas — it is in Denton, Texas.

THE TRUTH ABOUT T. REX

EVEN ONE OF THE BEST KNOWN DINOSAURS HAS KEPT SOME SECRETS. HERE IS WHAT PALAEOLOGISTS MOST WANT TO KNOW ABOUT THE FAMOUS TYRANT.

BY BRIAN SWITEK

In late 1905, newspaper reporters gushed over the bones of a prehistoric monster that palaeontologists had unearthed in the badlands of Montana. When *The New York Times* described the new 'Tyrant saurian', the paper declared it "the most formidable fighting animal of which there is any record whatever". In the century since, *Tyrannosaurus rex* has not loosened its grip on the imaginations of the public or palaeontologists.

Stretching more than 12 metres from snout to tail and sporting dozens of serrated teeth the size of rail spikes, the 66-million-year-old *T. rex* remains the ultimate example of a prehistoric predator — so much so that a media frenzy erupted this year over a paper debating

whether *T. rex* predominantly hunted or scavenged its meals¹. This infuriated many palaeontologists, who say the matter was resolved long ago by ample evidence showing that *T. rex* could take down prey and dismantle carrion. What particularly vexed researchers was that this non-issue overshadowed other, more important questions about *T. rex*.

The dinosaur's evolutionary origins, for example, are still a mystery. Researchers are eagerly trying to determine how these kings of the Cretaceous period (which spanned from 145 million to 66 million years ago) arose from a line of tiny dinosaurs during the Jurassic period (201 million to 145 million years ago). There is also considerable debate about what *T. rex* was like as a juvenile, and whether palaeontologists have spent decades mistaking its young for a separate species. Even the basic appearance of *T. rex* is in dispute: many researchers argue that the giant was covered in fluff or fuzz rather than scales. And then there is the vexing question of why *T. rex* had such a massive head and legs but relatively puny arms.

On the bright side, palaeontologists have material to work with. "We have lots of fossils of *T. rex*," says palaeontologist Stephen Brusatte of the University of Edinburgh, UK. "It's rare to have so many good fossils of one dinosaur, so we can actually ask questions about *T. rex* — such as how it grew, what it ate and how it moved — that we can't for other dinosaurs."

Here, *Nature* examines how palaeontologists are investigating these and other hot topics for the most charismatic of carnivores.

FUZZY ORIGINS

In the first few decades after palaeontologist Henry Fairfield Osborn named and described *T. rex*, researchers viewed this giant dinosaur as the culmination of a trend towards bigger predators. In this view, *T. rex* was seen as the descendent of *Allosaurus*, a 9-metre-long predator that lived more than 80 million years earlier. These and other massive carnivorous dinosaurs were lumped together in a categorical wastebasket called the Carnosauria, with *T. rex* as the last and biggest of the ferocious family. But palaeontologists tore up that evolutionary tree when they started using a more rigorous form of analysis called cladistics in the 1990s. They re-examined relationships between dinosaur groups and found that *T. rex* had its roots in a lineage of small, fuzzy creatures that lived in the shadow of *Allosaurus* and other predators during the Jurassic period.

The view that emerged placed *T. rex* and its close relatives — together known as tyrannosaurids — as the top twig on a broader evolutionary bush called the Tyrannosauroida, which emerged around 165 million years ago (see 'In the flesh'). Among the earliest

known members of this group was *Stokesosaurus clevelandi*, a bipedal carnivore 2–3 metres long that lived about 150 million years ago. Little is known about this creature, but evidence from other early tyrannosauroids suggests that *Stokesosaurus* had a long, low skull and slender arms. Early tyrannosauroids were small, agile predators, but their size placed them low in the pecking order during the Jurassic. "They were more lapdogs than top predators," says Brusatte.

The question for palaeontologists is how tyrannosaurs rose to power from such humble beginnings and why they took over as the apex predators in North America and Asia. At present, the key parts of this story are missing. There are relatively few dinosaur-rich rock formations from the period between 145 million and 90 million years ago, when tyrannosaurs apparently took over, so palaeontologists have yet to fully chart the communities that existed at the time. Shifts in sea level or climate could have triggered events that led to tyrannosaur dominance, Brusatte says, but he admits that such a connection is speculative. "We really need more fossils from this middle Cretaceous gap to help untangle this mystery."

In the past few years, researchers have started making headway in China, where rock formations record some segments of this key interval. In 2009, Peter Makovicky at the Field Museum in Chicago, Illinois, and his colleagues described a long-snouted tyrannosaur named *Xiongguanlong baimoensis* from rocks in western China dating to between 100 million and 125 million years ago². That animal reached about four metres long, a step up in size from the Jurassic tyrannosaurs. And, in 2012, Xu Xing of the Institute of Vertebrate Paleontology and Paleoanthropology in Beijing and his colleagues described a 9-metre-long tyrannosaur by the name of *Yutyrannus huali*³ from a similar time period (see *Nature* 489, 22–25; 2012).

This may be the crucial transition during which tyrannosaurs overlapped with allosaurs, before the latter faded out in the same habitats. In studies of rocks from northern China, Brusatte and his co-workers have found an allosaur five to six metres long named *Shaochilong maortuensis*, which lived about 90 million years ago⁴. "So it seems like both allosauroids and tyrannosauroids were around in Asia during this time, and had relatively similar sizes," he says. He hopes that further fossil discoveries will help to flesh out how and when tyrannosaurs took over as the top predator in their ecosystems.

ADOLESCENT ANGST

Just as the evolutionary origins of *T. rex* remain murky, so does its youth. In this case, the big debate centres on an creature called *Nanotyrannus lancensis*, a tyrannosaur found in the same North American deposits as *T. rex* that may have reached more than 6 metres in

T. REX ILLUSTRATION BY EMILY COOPER; FAMILY TREE FROM REF. 3

IN THE FLESH

Our picture of *Tyrannosaurus rex* has undergone several makeovers since the dinosaur was first described in 1905. Early reconstructions depicted a scaly beast that stood upright and dragged its tail on the ground, but recent research suggests the Cretaceous carnivore had a more agile horizontal posture and may have been covered in some sort of plumage.

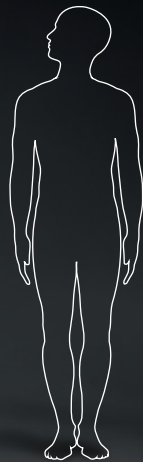
The small tyrannosaur known as *Nanotyrannus* (white skull) may have been a juvenile *T. rex* (skull outline).

Feathers on some close relatives of *T. rex* are more like fuzz than the plumage on birds.

If *T. rex* had a coat of proto-feathers, they may have served as a form of display.

Some researchers contend that *T. rex* and its kin had scaly skin.

Muscle scars on the arm bones suggest that the limbs were not vestigial.

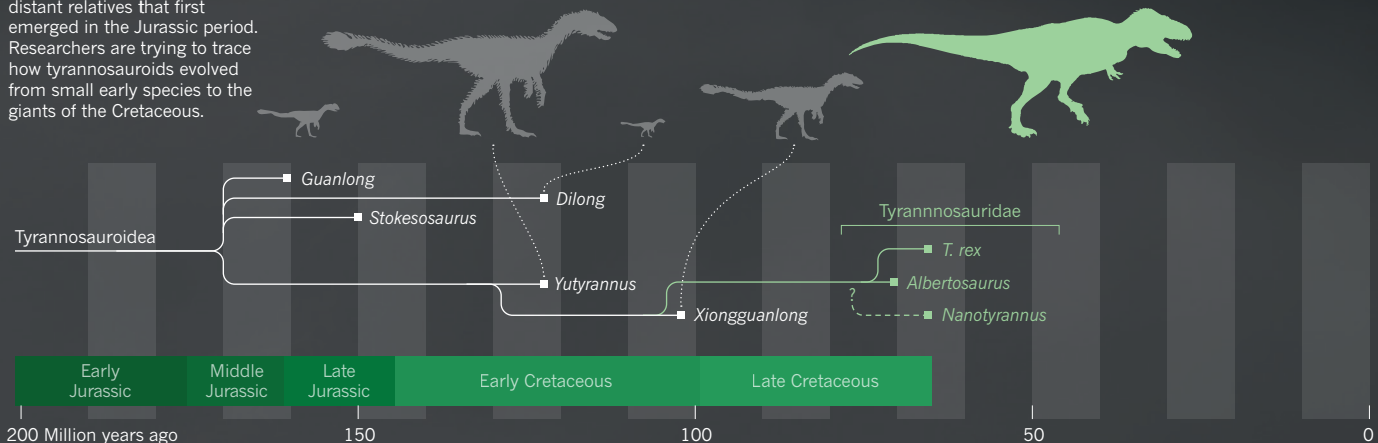


1905 reconstruction

T. rex was originally imagined with a reptilian, tail-dragging pose, but newer reconstructions make it a fleeter, more bird-like dinosaur.

TYRANNOSAUROID TREE

The tyrannosauroid superfamily includes Cretaceous tyrannosauroids, such as *T. rex*, and more distant relatives that first emerged in the Jurassic period. Researchers are trying to trace how tyrannosauroids evolved from small early species to the giants of the Cretaceous.



length. When it was first discovered, this creature was thought to be a separate species, but some researchers now argue that *Nanotyrannus* is actually just a juvenile *T. rex*.

According to Thomas Holtz Jr, a palaeontologist at the University of Maryland in College Park, *Nanotyrannus* specimens look remarkably like *T. rex*, and the differences between the two are similar to the differences between immature and mature individuals of other tyrannosaur species. The fact that all of the *Nanotyrannus* specimens seem to be juvenile animals and all of the specimens recognized as *T. rex* are subadults or adults, Holtz says, indicates that the two are truly one.

Lawrence Witmer, a palaeobiologist at Ohio University in Athens, is not so sure. In 2010, he and his colleague Ryan Ridgely studied computed-tomography scans of a skull from the Cleveland Museum of Natural History in Ohio that is the defining specimen, or holotype, of *N. lancensis*.

"We went into the project with the bias or assumption that the Cleveland skull was a

"IT IS BECOMING INCREASINGLY DIFFICULT TO REJECT A FUZZ-LESS TYRANNOSAURUS WITH A STRAIGHT FACE."

juvenile *T. rex*," Witmer says. But they found some unusual indentations in the brain case and sinuses, where air sacs filled the back of the skull in life⁵. These features are very different from those of *T. rex* and may identify the skull as belonging to a different species, says Witmer.

Team *Nanotyrannus* has no more vocal an advocate than Peter Larson, president of the Black Hills Institute of Geological Research, a company in Hill City, South Dakota, that collects, prepares and casts fossils. Larson argues that the teeth of *Nanotyrannus* are too finely serrated and closely packed to be those of a young *T. rex*. He also points to differences between the two species in the anatomy of the shoulder socket and the openings in the skull.

But some of these conclusions were gleaned from fossils not yet described in any publication, and scientists may never have a chance to study them. A skeleton that has been identified as a *Nanotyrannus* that could offer clues will be auctioned off next month in New York City. The hype generated by this specimen and its relevance to the *Nanotyrannus* debate has helped to drive up its price; estimates suggest that it may fetch up to US\$9 million. But most palaeontologists refuse to study such specimens unless they are placed in a reputable museum. A private buyer could rob researchers of that opportunity.

"The solution may reside in the tired plea for more fossils," Witmer says. For *Nanotyrannus* to have a shot at being a separate species,

palaeontologists would like to see one of two discoveries: a young tyrannosaur more similar to adult *T. rex* than any *Nanotyrannus* specimen, or an animal that is clearly an adult *Nanotyrannus* that is different from *T. rex*. But where an animal as charismatic as *T. rex* is concerned, it may be impossible for researchers to abandon long-held views and resolve decades of debate. "I'm not sure how much data it'll take to break us out of that," Witmer says.

A FLAP OVER FEATHERS

For generations, artists have depicted *T. rex* covered in scales, much like the modern-day reptiles to which it is only distantly related. But in the past two decades, researchers in China have found specimens from many dinosaur groups bearing feathers or a fuzzy coating. Some of these discoveries include species closely related to *T. rex*.

In 2004, Xu named *Dilong paradoxus* — a small, early tyrannosaur⁶. The fossil of this animal showed impressions of fibres around the tail, jaw and other body parts, suggesting the animal had a coat of 'dinofuzz'. The giant *Y. huali* from China also bore plumage³. The feathers on these tyrannosaurs were not like those of living birds, but simplified precursors. Xu suggests that the earliest feathered dinosaurs might have used their plumage for visual display. Later animals that were cloaked entirely in feathers might have relied on them for insulation. Because of the close evolutionary link between tyrannosaurs, he suggests that "*T. rex* might have had some kind of protofeathers".

Other researchers also favour the idea of feathered tyrannosaurs. "It is becoming increasingly difficult to reject a fuzz-less *Tyrannosaurus* with a straight face," Holtz says. That does not mean that *T. rex* looked like a Cretaceous chicken. Brusatte says it may have been covered in fairly inconspicuous hair-like fibres, like many other feathered dinosaurs.

As yet, no skin impressions have been found for *T. rex*, so researchers cannot say with certainty what kind of body covering it had. And some are not ready to abandon the more conventional view. Thomas Carr, a palaeontologist at Carthage College in Kenosha, Wisconsin, argues, for example, that unpublished fossils with skin impressions from close relatives of *T. rex* show scaly skin. These findings suggest that even though some earlier tyrannosaurids had feathers, the subgroup called tyrannosauridae (which includes *T. rex*), seems to have undergone an evolutionary reversal from fuzz to scales.

"There is no empirical evidence that tyrannosaurids had feathers," Carr says, "and artists have no business decking them out with plumage until the day comes when a tyrannosaurid is found with feathers."

This argument goes well beyond what the creatures looked like. Whether *T. rex* had feathers will influence how researchers

reconstruct the life of this dinosaur, from possible courtship behaviours to how it controlled its body temperature.

ARMS RACE

One of the biggest mysteries about *T. rex* has nagged palaeontologists for more than a century: what use did the giant have for arms so stubby that they could not even have reached its mouth? Early ideas, later discarded, suggested that the two-clawed arms helped *T. rex* to grip a partner during mating or to rise from repose. Later palaeontologists argued that the arms were vestigial — an idea beloved by cartoonists, who never tire of showing *T. rex* embarrassed by its useless, puny guns.

But research by palaeobiologist Sara Burch at Ohio University suggests that such jokes are unfair. She has studied the musculature of crocodylians as well as that of the only living members of the dinosaur line — birds. If the arms of *T. rex* had been vestigial, they would have lost the various anatomical landmarks that indicate muscle attachments, but the fossils "retain evidence of substantial musculature," she says.

But knowing that *T. rex* used its arms doesn't reveal what they were used for. To Carr, the arms were part of the dinosaur's arsenal. "Tyrannosaurids used their arms in the same way all theropods used their arms, for grasping and stabilizing objects" — namely prey, he says.

Holtz visualizes a less rigorous role for the forelimbs. On the basis of previous estimates of muscle strength, he argues that *T. rex* had weak arms. And because many tyrannosaurs have arms with healed fractures, he says, "their life habits could not require constant use of these arms". Holtz suggests that they were used primarily for display, perhaps during mating or competition — a possibility that seems more likely if these limbs were cloaked in feathers.

He and other palaeontologists plan to keep digging into the secrets of this superlative animal, one of the strongest ambassadors of the past in all of science. "Many aspects of *T. rex*, especially behavioural ones or physiological ones, are still unknown," Holtz says. But perhaps not forever. "As new methods of investigation are developed, we will have new avenues about their biology to explore." And as researchers do so, their views on the tyrant king will continue to evolve. ■

Brian Switek is a freelance writer in Salt Lake City, Utah.

1. DePalma, R. A., Burnham, D. A., Martin, L. D., Rothschild, B. M. & Larson, P. L. *Proc. Natl Acad. Sci. USA* <http://www.pnas.org/content/early/2013/07/10/1216534110> (2013).
2. Li, D., Norell, M. A., Gao, K.-Q., Smith, N. D. & Makovicky, P. J. *Proc. R. Soc. B* **277**, 183–190 (2010).
3. Xu, X. *et al. Nature* **484**, 92–95 (2012).
4. Brusatte, S. L., Chure, D. J., Benson, R. B. J. & Xu, X. *Zootaxa* **2334**, 1–46 (2010).
5. Witmer, L. M. & Ridgely, R. C. *Kirtlandia* **57**, 61–81 (2010).
6. Xu, X. *et al. Nature* **431**, 680–684 (2004).

By scanning blobs of brain activity, scientists may be able to decode people's thoughts, their dreams and even their intentions.

BY KERRI SMITH

Reading minds

“This is a manatee, but it doesn’t know what that is,” says Gallant, talking about the program as one might a recalcitrant student. They had trained the program, he explains, by showing it patterns of brain

Although companies are starting to pursue brain decoding for a few applications, such as market research and lie detection, scientists are far more interested in using this process to learn about the brain itself. Gallant's group and others are trying to find out what underlies those different brain patterns and want to work out the codes and algorithms the brain uses to make sense of the world around it. They hope that

these techniques can tell them about the basic principles governing brain organization and how it encodes memories, behaviour and emotion (see 'Decoding for dummies').

Applying their techniques beyond the encoding of pictures and movies will require a vast leap in complexity. "I don't do vision because it's the most interesting part of the brain," says Gallant. "I do it because it's the easiest part of the brain. It's the part of the brain I have a hope of solving before I'm dead." But in theory, he says, "you can do basically anything with this".

BEYOND BLOBOLOGY

Brain decoding took off about a decade ago¹, when neuroscientists realized that there was a lot of untapped information in the brain scans they were producing using functional magnetic resonance imaging (fMRI). That technique measures brain activity by identifying areas that are being fed oxygenated blood, which light up as coloured blobs in the scans. To analyse activity patterns, the brain is segmented into little boxes called voxels — the three-dimensional equivalent of pixels — and researchers typically look to see which voxels respond most strongly to a stimulus, such as seeing a face. By discarding data from the voxels that respond weakly, they conclude which areas are processing faces.

Decoding techniques interrogate more of the information in the brain scan. Rather than asking which brain regions respond most strongly to faces, they use both strong and weak responses to identify more subtle patterns of activity. Early studies of this sort proved, for example, that objects are encoded not just by one small very active area, but by a much more distributed array.

These recordings are fed into a 'pattern classifier', a computer algorithm that learns the patterns associated with each picture or concept. Once the program has seen enough samples, it can start to deduce what the person is looking at or thinking about. This goes beyond mapping blobs in the brain. Further attention to these patterns can take researchers from asking simple 'where in the brain' questions to testing hypotheses about the nature of psychological processes — asking questions about the strength and distribution of memories, for example, that have been wrangled over for years. Russell Poldrack, an fMRI specialist at the University of Texas at Austin, says that decoding allows researchers to test existing theories from psychology that predict how people's brains perform tasks. "There are lots of ways that go beyond blobology," he says.

In early studies^{1,2} scientists were able to show that they could get enough information from these patterns to tell what category of object someone was looking at — scissors, bottles and shoes, for example. "We were quite surprised it worked as well as it did," says Jim Haxby at Dartmouth College in New Hampshire, who led the first decoding study in 2001.

Soon after, two other teams independently used it to confirm fundamental principles of human brain organization. It was known from studies using electrodes implanted into monkey and cat brains that many visual areas react strongly to the orientation of edges, combining them to build pictures of the world. In the human brain, these edge-loving regions are too small to be seen with conventional fMRI techniques. But by applying decoding methods to fMRI data, John-Dylan Haynes and Geraint Rees, both at the time at University College London, and Yukiyasu Kamitani at ATR Computational Neuroscience Laboratories, in Kyoto, Japan, with Frank Tong, now at Vanderbilt University in Nashville, Tennessee, demonstrated in 2005 that pictures of edges also triggered very specific patterns of activity in humans^{3,4}. The researchers showed volunteers lines in various orientations — and the different voxel mosaics told the team which orientation the person was looking at.

Edges became complex pictures in 2008, when Gallant's team developed a decoder that could identify which of 120 pictures a subject was viewing — a much bigger challenge than inferring what general category an image belongs to, or deciphering edges. They then went a step further, developing a decoder that could produce primitive-looking movies of what the participant was viewing based on brain activity⁵.

From around 2006, researchers have been developing decoders for various tasks: for visual imagery, in which participants imagine a scene; for working memory, where they hold a fact or figure in mind; and for intention, often tested as the decision whether to add or subtract two numbers. The last is a harder problem than decoding the visual system says Haynes, now at the Bernstein Centre for Computational Neuroscience in Berlin, "There are so many different intentions — how do we categorize them?" Pictures can be grouped by colour or content, but the rules that govern intentions are not as easy to establish.

"Media reports have suggested that such techniques bring mind-reading 'from the realms of fantasy to fact'."

Gallant's lab has preliminary indications of just how difficult it will be. Using a first-person, combat-themed video game called *Counterstrike*, the researchers tried to see if they could decode an intention to go left or right, chase an enemy or fire a gun. They could just about decode an intention to move around; but everything else in the fMRI data was swamped by the signal from participants' emotions when they were being fired at or killed in the game. These signals — especially death, says Gallant — overrode any fine-grained information about intention.

The same is true for dreams. Kamitani and his team published their attempts at dream decoding in *Science* earlier this year⁶. They let participants fall asleep in the scanner and then woke them periodically, asking them to recall what they had seen. The team tried first to reconstruct the actual visual information in dreams, but eventually resorted to word categories. Their program was able to predict with 60% accuracy what categories of objects, such as cars, text, men or women, featured in people's dreams.

The subjective nature of dreaming makes it a challenge to extract further information, says Kamitani. "When I think of my dream contents, I have the feeling I'm seeing something," he says. But dreams may engage more than just the brain's visual realm, and involve areas for which it's harder to build reliable models.

REVERSE ENGINEERING

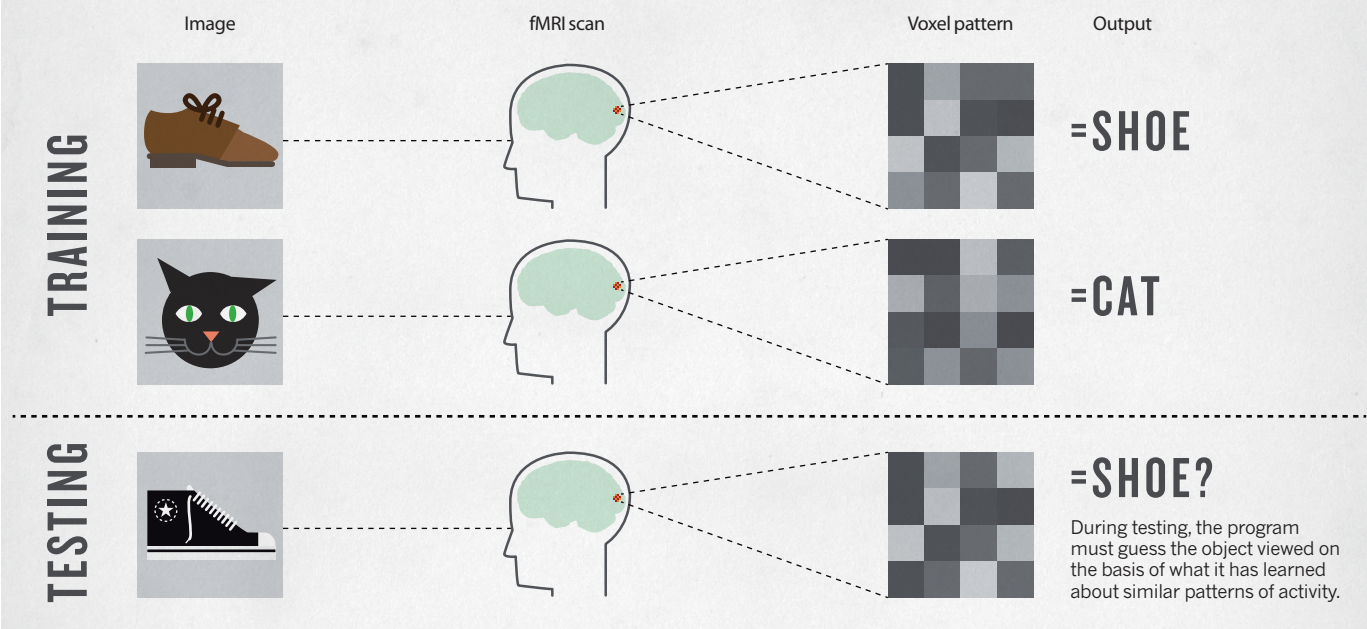
Decoding relies on the fact that correlations can be established between brain activity and the outside world. And simply identifying these correlations is sufficient if all you want to do, for example, is use a signal from the brain to command a robotic hand (see *Nature* **497**, 176–178; 2013). But Gallant and others want to do more; they want to work back to find out how the brain organizes and stores information in the first place — to crack the complex codes the brain uses.

That won't be easy, says Gallant. Each brain area takes information from a network of others and combines it, possibly changing the way it is represented. Neuroscientists must work out *post hoc* what kind of transformations take place at which points. Unlike other engineering projects, the brain was not put together using principles that necessarily make sense to human minds and mathematical models. "We're not designing the brain — the brain is given to us and we have to figure out how it works," says Gallant. "We don't really have any math for modelling these kinds of systems." Even if there were enough data available about the contents of each brain area, there probably would not be a ready set of equations to describe them, their relationships, and the ways they change over time.

➔ **NATURE.COM**
For a video of
this story go to:
go.nature.com/ocye5q

DECODING FOR DUMMIES

Scientists train a computer program by showing it brain-scan data associated with seeing certain images. Once it has built a database of activity patterns, it can be tested with images the participant hasn't necessarily seen before.



Computational neuroscientist Nikolaus Kriegeskorte at the MRC Cognition and Brain Sciences Unit in Cambridge, UK, says that even understanding how visual information is encoded is tricky — despite the visual system being the best-understood part of the brain (see *Nature* **502**, 156–158; 2013). “Vision is one of the hard problems of artificial intelligence. We thought it would be easier than playing chess or proving theorems,” he says. But there’s a lot to get to grips with: how bunches of neurons represent something like a face; how that information moves between areas in the visual system; and how the neural code representing a face changes as it does so. Building a model from the bottom up, neuron by neuron, is too complicated — “there’s not enough resources or time to do it this way”, says Kriegeskorte. So his team is comparing existing models of vision to brain data, to see what fits best.

REAL WORLD

Devising a decoding model that can generalize across brains, and even for the same brain across time, is a complex problem. Decoders are generally built on individual brains, unless they’re computing something relatively simple such as a binary choice — whether someone was looking at picture A or B. But several groups are now working on building one-size-fits-all models. “Everyone’s brain is a little bit different,” says Haxby, who is leading one such effort. At the moment, he says, “you just can’t line up these patterns of activity well enough”.

Standardization is likely to be necessary for many of the talked-about applications of brain decoding — those that would involve reading someone’s hidden or unconscious thoughts. And although such applications are not yet possible, companies are taking notice. Haynes says that he was recently approached by a representative from the car company Daimler asking whether one could decode hidden consumer preferences of test subjects for market research. In principle it could work, he says, but the current methods cannot work out which of, say, 30 different products someone likes best. Marketers, he says, should stick to what they know for now. “I’m pretty sure that with traditional market research techniques you’re going to be much better off”.

Companies looking to serve law enforcement have also taken notice. No Lie MRI in San Diego, California, for example, is using techniques related to decoding to claim that it can use a brain scan to distinguish a lie from a truth. Law scholar Hank Greely at Stanford University in

California, has written in the *Oxford Handbook of Neuroethics* (Oxford University Press, 2011) that the legal system could benefit from better ways of detecting lies, checking the reliability of memories, or even revealing the biases of jurors and judges. Some ethicists have argued that privacy laws should protect a person’s inner thoughts and desires as private, but Julian Savulescu, a neuroethicist at the University of Oxford, UK, sees no problem in principle with deploying decoding technologies. “People have a fear of it, but if it’s used in the right way it’s enormously liberating,” Brain data, he says, are no different from other types of evidence. “I don’t see why we should privilege people’s thoughts over their words,” he says.

Haynes has been working on a study in which participants tour several virtual-reality houses, and then have their brains scanned while they tour another selection. Preliminary results suggest that the team can identify which houses their subjects had been to before. The implication is that such a technique might reveal whether a suspect had visited the scene of a crime before. The results are not yet published, and Haynes is quick to point out the limitations to using such a technique in law enforcement. What if a person has been in the building, but doesn’t remember? Or what if they visited a week before the crime took place? Suspects may even be able to fool the scanner. “You don’t know how people react with countermeasures,” he says.

Other scientists also dismiss the implication that buried memories could be reliably uncovered through decoding. Apart from anything else, you need a 15-tonne, US\$3-million fMRI machine and a person willing to lie very still inside it and actively think secret thoughts. Even then, says Gallant, “just because the information is in someone’s head doesn’t mean it’s accurate”. Right now, psychologists have more reliable, cheaper ways of getting at people’s thoughts. “At the moment, the best way to find out what someone is going to do,” says Haynes, “is to ask them.” ■

Kerri Smith is senior audio editor for *Nature* in London.

1. Haxby, J. V. *et al. Science* **293**, 2425–2430 (2001).
2. Cox, D. D. & Savoy, R. L. *et al. NeuroImage* **19**, 261–270 (2003).
3. Haynes, J.-D. & Rees, G. *Nature Neurosci.* **8**, 686–691 (2005).
4. Kamitani, Y. & Tong, F. *Nature Neurosci.* **8**, 679–685 (2005).
5. Nishimoto, S. *et al. Curr. Biol.* **21**, 1641–1646 (2011).
6. Horikawa, T., Tamaki, M., Miyawaki, Y. & Kamitani, Y. *Science* **340**, 639–642 (2013).

COMMENT

AUTUMN BOOKS Craig Venter's vision of life as a machine **p.436**



AUTUMN BOOKS Physicists' moral struggle under the Nazis **p.441**

AUTUMN BOOKS Where did US mental-health-care policy go so wrong? **p.446**

OBITUARY Ronald Coase, whose economics inspired cap-and-trade, remembered **p.449**

JESSICA FARRER



Ecologist Jay Rotella (left) collects data in Antarctica on Weddell seals, with his students Glenn Stauffer and Thierry Chambert (right) in 2010.

The long shadow of the shutdown

Stalled Antarctic field work as a result of the US government shutdown has jeopardized early-career scientists and their projects, says **Gretchen E. Hofmann**.

For 13 seasons studying ocean-change biology in Antarctica, the familiar landmarks of Mount Erebus, Observation Hill and Castle Rock — visible even in a ground blizzard — have directed me back across the sea ice to the US research base, McMurdo Station. But when my postdoctoral researcher, Amanda Kelley, arrived on the sea ice of McMurdo Sound earlier this month, the direction of her research was not so clear.

As a polar research fellow, funded by the US National Science Foundation (NSF), Kelley was greeted with the news that her field season had been cancelled. On 1 October,

the US federal government shut down after Congress failed to agree on a budget for the next fiscal year. One week later, the ensuing lapse in funds for the US Antarctic Program (USAP; www.usap.gov) meant that McMurdo Station was switched to 'caretaker' status, meaning that activities focused solely on protecting personnel and property.

Even though a fiscal deal was reached in Washington DC on 16 October, ending this surreal shutdown, these few weeks of delay, of lost data and of instruments becoming ever more irretrievable, may have already irreparably damaged the season's research.

Furthermore, the current deal extends funding until only next January — before the end of the Antarctic research season. The consequences may be indelible, particularly for early-career scientists (see 'Out of season').

Projects on topics from ice-sheet dynamics to penguin ecology have been put at risk. Delaying work reduces the time that research in certain areas can be done; for example, soon, warmer weather will soften the sea ice and render it unsafe for travel.

I have a front-row seat to this stressful drama. The ocean-change biology research programme in McMurdo Sound that ▶

► I lead has been suspended indefinitely because of the shutdown. Kelley has gone to New Zealand and hopes to return later to McMurdo; the deployment to Antarctica of my graduate student, Lydia Kapsenberg, has been delayed indefinitely until further notice.

AGAINST THE CLOCK

Both these researchers' projects are time-sensitive. Kelley is examining responses to ocean acidification and warming in early-stage Antarctic sea urchins. Because urchins produce eggs in October, she needs to collect specimens immediately. In the first year of a two-year fellowship, her stipend will run out before her project is completed because of the delay.

Kapsenberg's research on ocean acidification requires pH measurements of Antarctic waters¹. Last November, we deployed an autonomous oceanographic sensor called a SeaFET^{2,3} at Cape Evans. That sensor, now bobbing under the sea ice just off shore from explorer Robert Falcon Scott's historic hut, holds the first winter data that we have ever recorded at our site. If we are unable to reach it before the sea ice melts, we will lose the data and perhaps, by next year, the sensor too.

Other McMurdo-based projects could be crippled if the research season cannot be resumed. Samantha Hansen, a geologist at the University of Alabama in Tuscaloosa, who is funded by a prestigious NSF CAREER award, needs to recover data from and service 15 instrumented stations installed last season to record global earthquakes and study the Transantarctic Mountains⁴. Costing tens of thousands of dollars each, the stations risk becoming non-functional or permanently buried beneath deep snow in the coming year if they are not maintained. Such a loss would affect her two graduate students, her Korean collaborators, and her linked NSF-sponsored education and outreach efforts.

Seal biologist Jay Rotella at Montana State University in Bozeman stands to lose more than most. His team has been counting and tagging Weddell seal pups in the McMurdo Sound area every year since 1968 (ref. 5). "If we miss this year," Rotella told me, "we will break the 31-year string of knowing every



Ecologists Gretchen Hofmann (left) and Paul Matson collect water samples in Antarctica.

pup that is born in the population and of recording complete reproductive histories for thousands of mothers." This, as bad luck would have it, is also the year that the current crew leader, PhD student Thierry Chambert, is scheduled to train his successor. Such handovers are crucial for the continuity of the project, which has trained generations of scientists.

Anne Todgham of San Francisco State University in California is the only project leader whose whole team made it to McMurdo Station before travel to Antarctica was cancelled by the shutdown. They will begin a new project to investigate the vulnerability of Antarctic fishes to climate change. "Time is of the essence," she wrote in an e-mail to me. "October is our window to collect eggs and juveniles." The early developmental stages of fishes are predicted to be the most affected by climate change⁶. And, as Todgham puts it: "Unfortunately, climate change does not stop for a government shutdown."

PERMANENT SCAR

Long-term data sets, many of which form foundations for studying climate change, stand to suffer serious damage. The NSF-funded McMurdo Long Term Ecological Research (LTER) group, for example, which has been monitoring ecosystems in the McMurdo Dry Valleys since 1993, could be put in jeopardy⁷. Because the lakes found in the valleys are terminal — nothing flows out of them — they are sentinels to climate

change. A lost season of water-column and sediment data could "compromise exponentially" our knowledge of how polar systems respond to climate change, says John Priscu, a microbiologist at Montana State University.

The shutdown means "a huge step backward in educating the next generation of polar scientists, and straying from the spirit of international collaboration that has been at the centre of Antarctic research," Priscu says. This year, he had hoped to work at McMurdo with a recent graduate from the Chinese Academy of Sciences.

Diana Wall, a soil ecologist from Colorado State University in Fort Collins, examines soil invertebrates such as the roundworm *Scottinema lindsayae*, considered the "toughest invertebrate in the Valleys"⁸. Its numbers have declined at McMurdo, affecting carbon turnover in Antarctic soil. She knows that gaps in records will cause problems for years to come. "I am really sad to think of the missing data and how hard it will be for the students, postdocs and me to explain any changes we miss."

With the end of the shutdown, there is now a possibility that some research at McMurdo could go ahead, assuming that logistical arrangements are still in place to support scientists. Todgham's team could still accomplish a great deal, and Rotella's group could tag this year's batch of seal pups if the sea ice remains open. Hansen's geologists, set to deploy in early November, might be able to retrieve their instruments. And the McMurdo LTER group could yet preserve this year's contribution to their time-series data. We are all crossing our fingers.

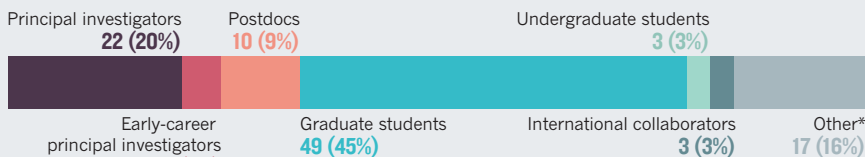
But the shutdown and its consequences are likely to leave a permanent scar on junior scientists, including my own, as thesis projects remain in jeopardy because it is yet unclear whether research will resume. Postdocs may eventually leave Antarctic science because the risks are too high to be borne by untenured researchers. Meanwhile, as principal investigators, we wait to hear the fate of our Antarctic research. ■

Gretchen E. Hofmann is professor of marine biology at the University of California, Santa Barbara, 93016 USA. e-mail: hofmann@lifesci.ucsb.edu

- Hofmann, G. E., Blanchette, C. A., Rivest, E. B. & Kapsenberg, L. *Oceanography* **26**, 140–148 (2013).
- Hofmann, G. E. *et al.* *PLoS ONE* **6**, e28983 (2011).
- Martz, T. R., Connery, J. G. & Johnson, K. S. *Limnol. Oceanogr. Methods* **8**, 172–184 (2010).
- Hansen, S. E. *et al.* *Geochim. Geophys. Geosyst.* **10**, Q08014 (2009).
- Chambert, T., Rotella, J. J. & Garrett, R. A. *Proc. R. Soc. B* **279**, 4532–4541 (2012).
- Munday, P. L., McCormick, M. I. & Nilsson, G. E. *J. Exp. Biol.* **215**, 3865–3873 (2012).
- Doran, P. T. *et al.* *Nature* **415**, 517–520 (2002).
- Barrett, J. E., Virginia, R. A., Wall, D. H. & Adams, B. J. *Glob. Change Biol.* **14**, 1734–1744 (2008).

OUT OF SEASON

The work of about a hundred US Antarctic Program personnel will be hindered by the suspension of the 2013–2014 field season due to the US government shutdown. More than half are early-career scientists.



*Educators and technicians. Data self-reported by personnel.



Set research priorities in a time of recession

Rigorous analyses are needed to establish the benefits of the knowledge economy, says former Irish government science adviser **Patrick Cunningham**.

Many governments say that they are using the current recession to refocus their public investment in science and technology. But after analysing countries' declarations of their research and development (R&D) funding and objectives to the Organisation for Economic Co-operation and Development (OECD) over the past decade, I have found that, in fact, not much has changed.

Nineteen of the 34 OECD member states have fully and consistently reported their civil R&D expenditure in the past two decades. Twelve of these have cut their public science budgets since 2007 (see go.nature.com/5dzkjp). Others have maintained modest growth. There have been exceptional annual funding increases in South Korea, a dramatic one-year stimulus in 2009 in the United States, and a European Union (EU) commitment to a 28% increase for its 2014–2020 budget. But research directions have remained the same.

I believe that all nations should use this time of change to improve the way that public funds are deployed in science. We need to learn from best practice at individual, institutional, corporate, national and international levels. To do so will require ongoing analysis of the facts, and a more rigorous scientific

approach to science policy.

Today, most OECD countries direct less than 1% of their tax revenues to R&D. This still amounts to substantial budgets under public control. The United States and the EU are responsible for half of the world's roughly US\$1,400-billion investment in R&D, despite being home to only 12% of its population.

Industry and businesses spend twice as much on R&D as governments do, split among thousands of enterprises. Despite this, the real driver of business innovation can be public expenditure: in the United States, the technological base of companies such as Apple, Intel, Google and much of the pharmaceutical industry is rooted in publicly funded research¹.

Governments vary widely in how they aim their R&D investment. The United States stands out as directing more than half of its budget to defence. By contrast, the EU spends 95% of its R&D investment on civil aims. Almost all other reporting countries had civil R&D fractions of more than 90% in 2011.

Civil R&D objectives as declared to the OECD fall into three classes: economic development, in sectors such as agriculture, industry and energy; specific public-good objectives, including health, environment, education, social and space programmes;

and non-oriented, or basic, research and general university funds. One might expect governments to favour economic impact in this time of austerity, but the OECD records show little shift in research spending focus.

Between 2006 and 2012, just one country out of 19 increased the amount it spent on economic objectives by more than 10%: Ireland (where I was the government's chief scientific adviser from 2007–2012) raised such spending by 13% to support innovation, growth and employment in the agriculture, food, marine and industrial sectors. And the country's business R&D expenditure rose by 43% between 2006 and 2010, although cause and effect are difficult to disentangle.

Most countries invested 20–30% of their science budget in economic development in 2011. South Korea, the highest such spender, targeted 50% as part of a purposeful and successful partnership between government and big business. Belgium and, with recently modernized economies, Finland and Ireland, spent just under 40% on economic development (see 'Civil spending shifts'). In the 1990s, Finland powered its way to economic recovery by increasing public investment in R&D, and, despite the recent travails of Finnish communications company Nokia, the country has weathered the latest recession better than most.

Countries with relatively low economic-development investment include the United States (11% in 2012) and the United Kingdom (8%), with large contributions to universities and defence.

The US and EU approaches to spending are radically different. Whereas the EU (taking into account each country's spending plus central European spending) directs more than half of its total civil R&D budget to non-oriented research and general university funding, most of the US civil R&D expenditure (73%) goes to health and environment programmes (see 'Different priorities').

In the United States, almost all public R&D funding comes directly from Washington DC, and this centralized system facilitates the scale, depth and continuity of programmes.

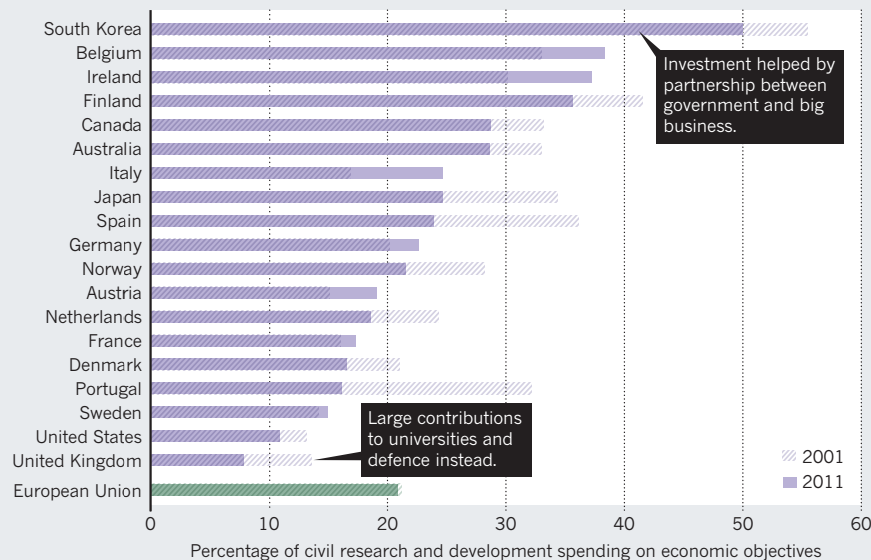
The more diffuse European funding structure can lead to duplication, but competition and diversity aid the spread of innovative ideas. Just 7% of EU research investment is channelled through Brussels, although this might rise to 10% under Horizon 2020, the next EU research and innovation funding cycle that will run from 2014 to 2020.

The overall level of US and EU spending on R&D has changed little in the past decade; it is still too early to judge the impact of the \$20-billion spike in US-government

“Industry and businesses spend twice as much on R&D as governments.”

CIVIL SPENDING SHIFTS

Countries including South Korea, Ireland and Finland have focused their civil research and development investment on economic development, in areas such as energy and industry.



R&D funds (a 14% hike) in 2009. In 2000, the ambition of the EU Lisbon Strategy to spend 3% of gross domestic product (GDP) by 2010 on public and private R&D combined was achieved by only three countries: Finland, Sweden and Denmark. For the EU as a whole, the figure is under 2%. And at just under 0.7%, public investment is still well below the Lisbon target of 1%.

Evaluating the impacts of R&D is challenging because they might not be felt until many years after the publication of research, and credit is difficult to apportion. The main challenges are to clarify the timescales involved and to quantify the trade-offs and synergies among inputs, outputs and interactions with parallel developments in other countries and in the business sector.

Better models and metrics need to be developed to measure the inputs, outputs and progress of the knowledge economy. The US Science of Science Policy initiative² (see scienceofsciencepolicy.net), which was proposed in 2005 by physicist and presidential science adviser John Marburger³, has made a start. Some 150 research contracts have been awarded to analyse the social and administrative structures of research programmes and their links with sectors of society. But the US focus is on its centralized structures.

PICK PRIORITIES

Europe, where the flows of many smaller national investments need to be understood, lags behind in science-policy analysis. Data collected by the OECD and Eurostat have informed cross-country studies, such as the Innovation Union Scoreboard⁴ that assembles 25 indicators into an innovation index. Countries such as China, Japan, South Korea

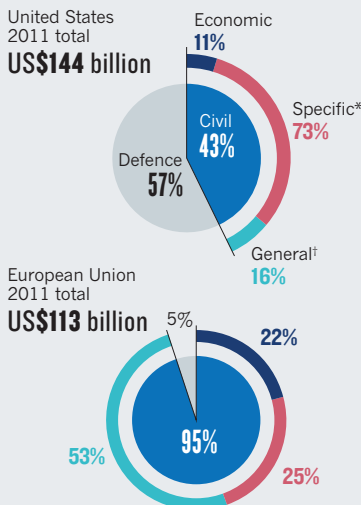
and Taiwan also lack substantial science-policy analysis programmes.

In this time of recession, when taxpayers are asked to invest their hard-earned money for the public good, all governments need to reassess the aims of their R&D budgets. Each nation must decide its own priorities; the experiences of Ireland and Finland suggest that there is much to be gained by investing explicitly for economic development — the benefits might be evident within a few years. The merits of defence research require debate.

The level of R&D funding needs to be raised across the board. EU governments

DIFFERENT PRIORITIES

Public research and development budgets are skewed towards defence in the United States and mainly towards civil programmes in Europe.



*Health, environment, education, social and space programmes
† University funds and non-oriented research

should recommit to the Lisbon Strategy goal and boost their public funding of R&D to 1% of GDP as soon as possible. Private investment should follow with encouragement, as in Ireland.

Better economic models are needed to understand the impacts of investments in different areas. These could follow the framework set out in two World Bank reports^{5,6} that consider natural resources, produced goods and services, and intangible social and intellectual value capital analogously to economist Adam Smith's 'land, labour and capital'. The first two are readily measured; the last is hard to evaluate but constitutes most of the wealth in developed societies.

In the meantime, GDP growth is a reasonable aim for R&D investment. Although it will not deliver all of the benefits that society desires, it correlates closely with broader measures such as the Human Development Index and Satisfaction with Life Index. GDP is thus not an end in itself, but an enabler of multiple end points.

To understand linkages between R&D investment and societal benefit, the 'science of science policy' field must be developed. The EU's 28 national programmes deserve attention because they constitute a rolling experiment in building knowledge economies. A series of workshops and joint research calls is needed to bring scientists and economists together to study the effects.

Europe will benefit from pooling its diverse experiences to get better value from the more than 90% of its R&D spend that is locked into national budgets. By strengthening links between researchers and institutions, perhaps through the EU Joint Programming Initiative, EU countries will gain more from Horizon 2020 than the financial contributions they make.

Most of the public science budget is invested in people, and most research is conducted by young scientists who move on to deploy their knowledge and skills throughout the economy. Governments must acknowledge that R&D is the driver of future welfare, security and prosperity. ■

Patrick Cunningham is professor of animal genetics at Trinity College Dublin, Ireland. From 2007 to 2012 he was chief scientific adviser to the Irish government.
e-mail: epcnngm@tcd.ie

1. Mazzucato, M. *The Entrepreneurial State* (Anthem, 2013).
2. Lane, J. *Nature* **464**, 488–489 (2010).
3. Marburger, J. H. *Science* **308**, 1087 (2005).
4. European Commission. *Innovation Union Scoreboard 2013* (EC, 2013); available at <http://go.nature.com/sxvo7d>.
5. World Bank. *Where is the Wealth of Nations?* (World Bank, 2006); available at <http://go.nature.com/xr2iju>.
6. World Bank. *The Changing Wealth of Nations* (World Bank, 2011); available at <http://go.nature.com/n3j3sj>.

AUTUMN BOOKS



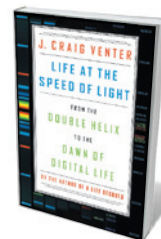
GENETICS

The genetic watchmaker

Nathaniel Comfort assesses Craig Venter's vision of nature-as-machine.

For centuries, the metaphor of nature-as-machine served as evidence of design in the Universe — and the existence of a Divine Machinist. William Paley's famous 1802 image of the watch and watchmaker prodded Charles Darwin towards his naturalistic theory of evolution. Machine metaphors remain ubiquitous in modern biology, but today, mechanisms such as 'clocks', 'signalling', 'transport', 'molecular hinges' and enzymatic 'locks and keys' are invoked reflexively — almost automatically — and each gives tacit testimony against vitalism, the belief in an ineffable life force.

In his characteristically brash, lively book, biologist Craig Venter gives us nature-as-computer, rigidly deterministic and controlled by the central program of DNA. Splicing an account of his own genomics research onto a historical trajectory, *Life at the Speed of Light* is a story about science accelerating towards total mastery of the living world. "This new understanding of life, and the recent advance in our ability to manipulate it," he writes, is leading us into "an era of biological design. Human-kind is about to enter a new phase of evolution." For Venter, life is an information



Life at the Speed of Light: From the Double Helix to the Dawn of Digital Life
J. CRAIG VENTER
Viking Adult: 2013.

system, intrinsically digital and hence as manipulable as software. His vision is to code, debug and compile synthetic organisms that will make us and our environment healthier, more harmonious, better.



ILLUSTRATIONS BY MARTIN O'NEILL; PHOTO: JESSICA RINALDI/REUTERS/CORBIS

Venter first traces the slow, steady march of the engineering ideal in biology. He examines thinkers such as the philosopher Francis Bacon, whose *New Atlantis* (1623) portrayed a scientific Utopia in which man established “dominion over Nature”, and the physiologist Jacques Loeb, who said in 1905 that “control and nothing else is the aim of biology”. So eager is Venter to exterminate vitalism from science that he treats the concept of emergent properties — the notion that the whole can be more than the sum of its parts — as vitalistic. But emergence need not require some spooky mystical glue; it can be accommodated by ordinary physics and chemistry. I can see, though, why it makes Venter uncomfortable: it introduces indeterminacy.

NATURE.COM
For a review of
Craig Venter's
autobiography, see:
go.nature.com/8fr2s8

Venter then gives the standard account of the mid-twentieth-century rise of the molecular view of

life, up through the double helix, the genetic code and the tools for sequencing and synthesizing DNA. He lingers affectionately on the contributions of his friend and collaborator Hamilton Smith to early recombinant-DNA research, which ties the history to the memoir.

By the 1980s, two information sciences were firmly established: molecular genetics, with its jargon rich in metaphors of text and information; and computer science. Venter's innovations have involved merging them. He has a talent for thinking algorithmically about DNA, for bold approaches that use huge computing power to gain dominion over the genetic material, and he has a keen business sense. His method of expressed sequence tagging, for example, identified thousands of genes and triggered a controversy when his employer (the US National Institutes of Health) attempted to patent them. Venter raised venture capital and went private, going head-to-head with his former employer in the race to sequence the human genome, and vastly accelerating the Human Genome

Project with his new method of ‘shotgun’ sequencing. To pull off this feat, Venter's company Celera assembled the largest and most powerful computer in the civilian world, which could handle 80 terabytes of data using 64 gigabytes of RAM — more than 1,000 times that of a high-end personal machine at the time.

Venter continued to ask engineering questions, such as, “what is the minimal genome that can support life?”. Turning to nature for a model, he selected one of the smallest known genomes, that of the virus Φ X174, which was sequenced by Fred Sanger in 1977. Venter and his team then synthesized that sequence, chunk by oligomeric chunk, and stitched the pieces back together to make a complete genome. In 2008, they repeated this feat with *Mycoplasma genitalium*, a bacterium that has the smallest genome of any organism that can be cultured. With a flourish, they branded the first “synthetic genome” by including a sequence that spelt out the names of the collaborators — ‘Venter Institute’ and ‘Synthetic Genomics’ — like a “watermark” on a document, as Venter says.

Their next step was to take the natural genome of *Mycoplasma mycoides* and insert it into the de-genomed husk of the related species *M. capricolum*. Venter calls this “converting one species into another”. Finally, the Venter team repeated the trick with an *M. mycoides* genome that came out of a DNA sequencer. This time, along with the brand names, they stuffed a DNA-coded message into their cellular bottle that said, in effect, ‘If found, please contact ...’, providing the organism with its own e-mail address. Venter called this a “synthetic cell”

and dubbed it a new species: “*M. mycoides* JCVI-syn 1.0”.

So much for the book's coding region. After sketching some of his ongoing projects, Venter speculates on the future. He gives us his own New Atlantis, a secular genotopia in which novel DNA sequence will be synthesized to specification, “teleported” at light speed and printed out on biological three-dimensional printers. Fans of *Star Trek*, however, know that teleporters do not leave the original behind. More aptly, DNA would move into the Cloud,

infinitely copyable from anywhere. Novel synthetic life forms, Venter writes, could help to solve some of society's most pressing problems. Climate change? There would be bio-apps for that, such as engineered algal biofuels. Famine? Drought? Ditto. It is biology for the Google set: unsentimental, joyfully technophilic and boundlessly optimistic.

There is geeky cool in this view of life, but little grandeur. Venter won't brook the complexities of Darwin's tangled banks: to make his claims, as he admits, he must clean up messy terms such as ‘life’, ‘organism’, ‘species’ and ‘teleportation’ for the laboratory. In effect, biology becomes what the J. Craig Venter Institute produces. The machine in the metaphor now is the JCVI itself. And the watchmaker, of course, is Venter. ■

Nathaniel Comfort is professor of the history of medicine at Johns Hopkins University in Baltimore, Maryland. His most recent book is *The Science of Human Perfection: How Genes Became the Heart of American Medicine*.
e-mail: nccomfort@gmail.com

**VENTER'S
VISION
IS TO CODE, DEBUG AND
COMPILE SYNTHETIC
ORGANISMS THAT
WILL MAKE US AND
OUR ENVIRONMENT
HEALTHIER,
MORE
HARMONIOUS,
BETTER.**



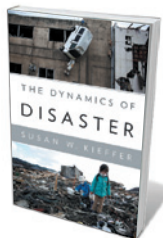
DISASTER MANAGEMENT

Preparing for the worst

A study on natural disasters puts fizz into the physics, finds **Roger Bilham**.

Susan Kieffer's *The Dynamics of Disaster* will attract readers much as a road accident slows a passing motorist. But those seeking Hollywood-style gore and fright are in for an education. Kieffer's geophysical study is much more than a litany of bad things happening to people who are in the wrong place at the wrong time. It delves into the physics responsible for many of the extreme events that society finds inconvenient, and offers hope that, rather than meekly accepting the rubbish that nature throws at us, we can attempt a societal fix.

Do not be put off by the rather dull introductory chapter, in which Kieffer dispenses some necessary definitions of disaster and places her book in context. Catastrophic surprises (such as earthquakes) or insidious change (global warming) have an obvious common denominator: in a world without people, disasters do not exist. One is reminded of graffiti scribbled in the 1960s on a wall in Cambridge, UK — "Hair needs a comb" — beneath which an undoubtedly long-haired student had scrawled "but not as much as a comb needs hair".



The Dynamics of Disaster

SUSAN W. KIEFFER
W. W. Norton: 2013.

quakes, volcanoes, cyclones, landslides and tsunamis, but lurking within these pages are some less familiar oddities — quick clay, lateral blasts, explanations of Mach numbers and rotating volcanic plumes.

It is Kieffer's gung-ho approach to the underlying mechanisms of all these extreme events that really makes this book interesting. Throughout, she invokes analogies and personal experiences to explain some of the more elusive concepts, and many that are less so. Her well-meaning comparisons are sometimes a bit odd, for instance: a tsunami taller than any mountain in Minnesota; "to

The book's theme is that disasters are characterized by a change of state from normal to briefly abnormal. What is intriguing is the breadth of extreme geological events that Kieffer invokes and explains, given this basic view of Earth's processes. We expect to read about earth-

sprint eighty-six storeys up to escape this wave"; "landslides are like robbers"; and "waves are rather like teenagers". But, as literary tools in the hands of a clever scientist, they do force the reader to grapple with the sometimes prodigious numbers involved.

Some will find the exuberant subheadings vexing. But at least lines such as "Shake, bake, zap, and glow" will grab the attention of politicians (and undergraduates who are poised to start texting in class), drawing them into the easy authority with which she explains the atmospheric features known as Hadley cells and the complexities of tsunami generation.

Kieffer is at her best when describing the fluid dynamics of the climate, atmosphere and oceans — this section is a good read for a solid-Earth scientist who wonders what all the fuss is about above ground. For example, I found her discussion of rogue waves (which may be responsible for the loss of 30 ships each year) surprisingly interesting.

On earthquakes, her explanations are a trifle misleading. Although liquefaction certainly contributes to the damage caused by earthquakes (such as those in Christchurch, New Zealand, in 2010 and 2011), its onset

MARTIN O'NEILL

**NEW IN
PAPERBACK**

*Highlights of this
season's releases*



The Future: Six Drivers of Global Change

Al Gore (Random House, 2013)

Former US vice-president and prominent voice in climate politics Al Gore tackles six areas of rapid change that are transforming our world — from the Internet and environmental crises to globalization and population growth. Gore's analyses of the scientific, political and economic aspects of each are thorough and compelling as he works towards a cautiously optimistic synthesis. (See Barbara Kiser's review: *Nature* **494**, 429; 2013.)

is not instantaneous but follows minutes after the earthquake. Liquefaction in Haiti's earthquake disaster of 2010 was responsible for few fatalities, with most of the damage occurring on bedrock.

The occasional jibes at the insensitivity and ignorance of myopic politicians will raise a cheer from many readers, as will Kieffer's championing of the precautionary principle. Simply stated, it is not up to the suffering world to prove that it is suffering. More precisely, if a government sanctions actions that may be harmful to our environment, it is up to the perpetrators to prove that their deeds are harmless. The principle applies well to profitable corporations. But how does it apply to unregulated deforestation by the world's poor, or to those who drive their cars to work?

At the end of each chapter, Kieffer explores the societal implications of the disasters, the threads of which she gathers in her concluding chapter. For instance, the double disaster in L'Aquila, Italy (the fatal earthquake of 2009 and its unexpected legal consequences), raises an important issue all scientists must face — how to describe uncertainty to a public that wants a black-and-white view of the future. In Italy, government representatives have chosen the moral high ground in condemning the absence of a clearly stated probabilistic assessment of potential future seismicity. Kieffer rightly views the L'Aquila process as a wake-up call for improving tools for characterizing future disasters. In a post-Fukushima world, we cannot afford to suppress an honest discussion of low-probability extreme events. But assessing what constitutes an acceptable risk to society is currently something that scientists and present societal structures are ill-equipped to handle.

Anyone interested in the processes that underlie catastrophic events within Earth will welcome this book, part riveting and all informative. We cannot prevent disasters, but with a little bit of foresight and a lot of common sense, we can reduce their impact on our growing population. Give a copy to your local politician! ■

Roger Bilham is a professor of geology at the University of Colorado in Boulder. He has published more than 200 articles on aspects of earthquakes and their effects on society. e-mail: roger.bilham@colorado.edu



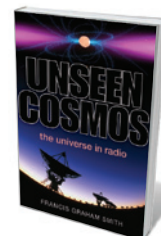
RADIO ASTRONOMY

Finger on the pulsar

Bernie Fanaroff probes a study on how radio telescopes have opened up our understanding of the Universe.

Francis Graham-Smith's *Unseen Cosmos* sets out the unique role of radio telescopes and observations at radio wavelengths in transforming our understanding of the Universe. The former UK Astronomer Royal describes the many important discoveries in radio astronomy and the techniques that made them possible. It is an extraordinary tour, from the rotating ultra-dense neutron stars known as pulsars and the cosmic microwave background left over from the Big Bang to powerful, distant radio-wave-emitting galaxies and the radio emission from molecules in galactic regions where stars are born.

Astronomy today is a multi-wavelength discipline. Observing astronomical objects and even the structure of the Universe at wavelengths from radio waves to gamma rays allows us to see different processes and often different parts of these objects.

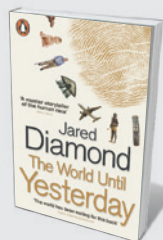


Unseen Cosmos: The Universe in Radio

FRANCIS GRAHAM-SMITH
Oxford University Press: 2013.

Observations in the infrared reveal cool galactic gas and dust; in the ultraviolet, hot young stars. At radio wavelengths, we spot neutral hydrogen gas and its motion, as well as synchrotron radiation (from electrons moving in a magnetic field at close to the speed of light) in galactic or intergalactic magnetic fields. X-ray telescopes detect very hot gas in and between galaxies, and optical wavelengths reveal the light from stars and ionized gas clouds. All of these data must be combined for a full understanding of objects. ▶

MARTIN O'NEILL; BASED ON A PHOTO BY NEILLD/BIGSTOCK



The World Until Yesterday

Jared Diamond (Penguin, 2013)

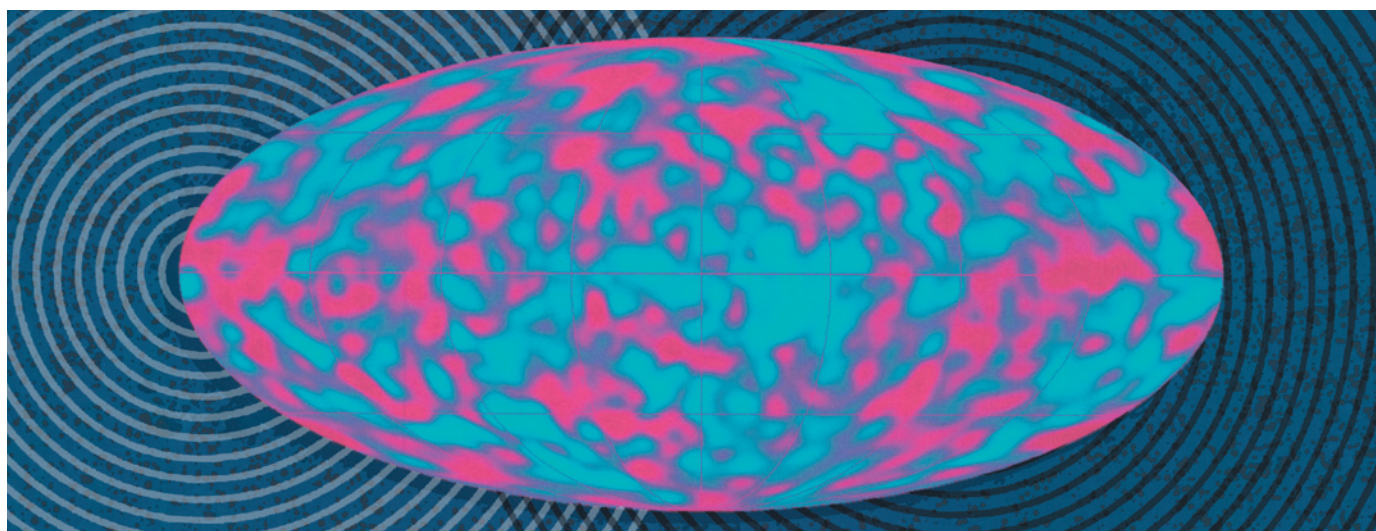
The cultural gap between traditional societies and the West is a rich seam for anthropologist Jared Diamond. Here, he explores what indigenous cultures can teach the West in areas from childcare to dispute resolution. (See Monique Borgerhoff Mulder's review: *Nature* **493**, 477–478; 2013.)



The Universe Within: The Deep History of the Human Body

Neil Shubin (Vintage, 2013)

Palaeontologist Neil Shubin unpicks the intertwined evolution of Earth and life, finding intriguing links, for example, between continental break-up and mammalian evolution. (See Birger Schmitz's review: *Nature* **493**, 25; 2013.)



MARTIN O'NEILL; NASA/COBE SCIENCE WORKING GROUP

► Multi-wavelength observation is also needed because many astronomical phenomena are now known to be intimately linked. The evolution of galaxies and clusters of galaxies is a good example: there are complex, still little-understood relationships between phenomena such as radiation and jets from active galactic nuclei (AGNs, regions at galactic centres that emit vast amounts of energy, powered by supermassive black holes), accretion of gas, star formation and galaxy mergers. Observing galaxies at different epochs, stages of development and wavelengths is helping to clarify how energy is transferred between AGNs and the gas in and between galaxies, and how this affects the rate of star formation.

Against these new trends in astronomy, it is easy to forget radio astronomy's special role over the past 80 years. Graham-Smith reminds us that the existence of the Big Bang was confirmed initially by counting distant radio galaxies and radio quasars — remote, extremely luminous AGNs — and then by the discovery of the cosmic microwave background. He describes the beautiful experiments that measured the irregularities in this radiation and how they have transformed cosmology from a science based at least in part on aesthetics to one in which key parameters have been determined to an extraordinary level of precision. He details the discovery of pulsars by Jocelyn Bell and Tony Hewish and the extreme physics of

these stars. The use of rapidly rotating pulsars as clocks has allowed astronomers to probe physics in very strong gravitational fields and has repeatedly confirmed the predictions of Einstein's General Theory of Relativity.

The new radio telescopes — such as the Square Kilometre Array (SKA) to be built in southern Africa and Australia, which will be the largest ever — also open up big possibilities. We could discover how the Universe was re-ionized by the first stars and/or quasars, detect the gravitational waves predicted by Einstein and possibly even detect extraterrestrial intelligence. The SKA will be sensitive enough to see ambient radio emission (the equivalent of airport radar) from habitable planets orbiting stars in our vicinity, and is by far the most likely way to find ET.

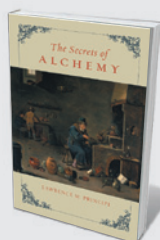
The first radio-astronomical observations were carried out by Karl Jansky and Grote Reber in the 1930s, but the key technological advances took place after the Second World War. Astronomers such as Martin Ryle, John Bolton, Bernard Lovell and Graham-Smith himself were amazingly innovative in designing and developing new instruments, such as radio interferometers. I was lucky to be a research student at the University of Cambridge, UK, from 1970 to 1974, with access to the One-Mile Telescope and 5-km Array. This was a unique opportunity — everything observed was new, exciting and publishable. *Unseen Cosmos* describes this history. And the tradition of innovation has persisted: the

technology challenges in designing and building the SKA are immense. They range from wide-field and wide-bandwidth receivers to innovative algorithms for calibrating and making images from observations. The vast data output will stretch researchers' capacity.

Although much of the history has been told before, I found *Unseen Cosmos* interesting and informative. Combining history with explanations of particular topics and their contemporary development has its limitations, however. And like most books that try to describe very complex physics in a simple way, this book succeeds in some places and not in others. I found the description of pulsars lengthy but hard to understand. I would also have welcomed more on current developments and what capabilities will be provided by the new radio telescopes, such as the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile and the SKA.

Because radio astronomy is developing rapidly, it is perhaps safer to write a book that includes a large dollop of history than to write one that could quickly become dated. Nonetheless, this book is a useful reminder of why we want to build huge, technically challenging and expensive radio telescopes like the SKA. ■

Bernie Fanaroff is the project director of South Africa's SKA project. He was Deputy Director General of President Nelson Mandela's Presidency.
e-mail: bfanaroff@ska.ac.za



The Secrets of Alchemy

Lawrence Principe (University of Chicago Press, 2013)
The practice of alchemy overlapped with the birth of chemistry, reveals Lawrence Principe in this magisterial study. He traces its trajectory from ancient Egypt through its development in the Islamic world, Latin Europe and beyond. (See Jennifer Rampling's review: *Nature* **491**, 38; 2012.)



Round About the Earth: Circumnavigation from Magellan to Orbit

Joyce E. Chaplin (Simon & Schuster, 2013)
The ultimate round trip, circumnavigation has seduced scientists and explorers for five centuries. This riveting history covers sea, land, air, space, and transport from feet to Sputnik. (See Andrew Robinson's review: *Nature* **491**, 39; 2012.)

PHYSICS

Science under the Nazis

Robert P. Crease applauds the story of three great physicists who struggled to maintain their integrity during the Third Reich.

Beware! This book is not what it seems. The subtitle suggests a black-and-white tale of good and evil, to be read in detached comfort from high moral ground. Instead, science writer Philip Ball delivers an ambiguous yet moving saga of well-intentioned people compelled to act in “the grey zone between complicity and resistance”. Its disturbing implications will leave attentive readers uneasy.

Ball follows the lives of three Nobel laureates under the Third Reich: Max Planck, Peter Debye and Werner Heisenberg. Planck was a humble member of the German intellectual elite who devoted himself to state service and, as head of the Kaiser Wilhelm Society (KWS), which promoted the natural sciences in Germany, was the titular representative of German science. Debye, a political and scientific pragmatist, was born in Maastricht, the Netherlands, but obtained nearly all his scientific training in Germany and professed himself culturally German. In 1934, he became director of the Kaiser Wilhelm Institute for Physics in Berlin. The ambitious and arrogant Heisenberg often acted as though he was the personification of German physics.

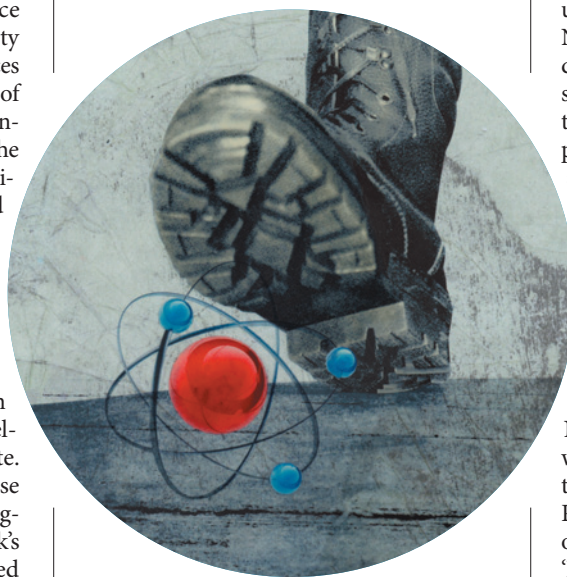
The disturbing saga begins in 1933, when Adolf Hitler was appointed Reich chancellor, paving the way for a totalitarian state. The Nazis increasingly forced Planck to use the KWS for political ends, such as by purging Jewish members — including Planck's friend Albert Einstein. Debye was coerced into a similar situation at the Kaiser Wilhelm Institute until he left for the United States in 1939. Heisenberg was a principal architect of the German atomic-bomb project. Ball traces how the Nazis ruthlessly exploited these and other scientists by preying on personal weaknesses and political naivety — citing “Debye's occasional self-interest and limited



Serving the Reich: The Struggle for the Soul of Physics Under Hitler
PHILIP BALL

Bodley Head: 2013.

moral engagement, Heisenberg's insecurity and egotism, Planck's prevarication and misconceived notion of duty” — to wheedle and compel them into actions that now look



“disturbingly compliant” at best, and utterly immoral at worst.

Yet Ball does an outstanding service by reminding us how powerful and sometimes confusing the pressures were, and how it was not implausible to think that scientists could and should stay ‘above politics’. Nazi

tyranny and genocide were unprecedented, yet aspects of their programme seemed progressive, including their welfare and healthcare policies, and efforts to eliminate class differences. Moreover, the Nazi party was not monolithic, but comprised rival factions competing for Hitler's favours throughout the Reich, and was plagued by incompetent leaders and an inept bureaucracy. Many observers, inside Germany and out, including the three physicists in question, assumed not unreasonably that the Nazis would be forced to moderate their behaviour or lose power.

For these and other reasons, Ball writes, understanding moral behaviour under the Nazis is not “a matter of simply collating the documentary evidence and totting up episodes of compliance or resistance”. He contends that we have to mine the ambiguous phrases and equivocal actions of scientists, and explore their inability to fathom their own motivations to reach a deeper understanding of their characters in a burgeoning atmosphere of paranoia and brutality.

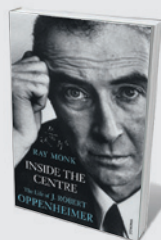
Serving the Reich is packed with dramatic, moving and even comical moments. One is the harrowing story of Austrian-Jewish scientist Lise Meitner's escape from Germany in 1938. Her Nazi neighbour alerted the authorities, but word failed to reach the border patrol in time. More touching is an anecdote about Planck presiding at an official function and only managing to utter the abhorrent phrase ‘Heil Hitler’ on his third attempt. And Debye, anticipating that the Nazis would refuse to let him rename a science institute after Planck, carved Planck's name into the stone above the entrance. When ordered to remove it, Debye covered it with a wooden plank (the pun also works in German).

Ball recounts Heisenberg's famous visit to occupied Copenhagen in September 1941, ►



Mirror Earth: The Search for Our Planet's Twin
Michael D. Lemonick (Bloomsbury, 2013)

Science writer Michael Lemonick explores astronomers' interest in sister worlds. Focusing on NASA's Kepler space telescope, this book is studded with in-depth portraits of exoplaneteers such as David Charbonneau, hunter of super-Earths. (See Sara Seager's review: *Nature* **490**, 479; 2012.)



Inside the Centre: The Life of J. Robert Oppenheimer
Ray Monk (Vintage, 2013)

This testimony to the triumphs and foibles of J. Robert Oppenheimer is illuminating. Ray Monk follows the physicist from adolescence to his role in the construction of the first atomic bomb. (See Istvan Hargittai's review: *Nature* **491**, 670; 2012.)

► where he annoyed fellow scientists with his grandiosity. His self-delusion persisted: early in 1945, after a special allied mission had raced across a collapsing Germany to apprehend him, Heisenberg arrogantly assumed that he held a powerful bargaining position and evidently failed to grasp that he was a prisoner. When he heard that the United States had dropped an atomic bomb on Hiroshima, at first he refused to believe it, claiming that some “dilettante” American had to be bluffing.

Although such scenes make *Serving the Reich* a page-turner, Ball keeps the moral and existential ambiguities at the forefront. He lets us see that for many scientists, to abandon one's work and post — especially during such a crisis — would seem “a dereliction of duty, not a moral act of protest”. And defying the Nazis was not always an act of rebellion: Planck's insistence on holding a memorial in 1935 on the first anniversary of the Jewish scientist Fritz Haber's death was less a protest against anti-Semitism than an honour extended to a deceased, esteemed colleague.

But Ball has no sympathy for journalists who have bought scientists' self-serving apologies or condemned the scientists on the basis of cherry-picked evidence. Dutch journalist Sybe Rispens's 2006 accusation that Debye was a Nazi sympathizer, for instance, led the University of Utrecht in the Netherlands to drop the physicist's name from its nanomaterials institute.

Ball insists that, rather than simplistically condemning or absolving the German scientists, we should look at their moral behaviour as a perpetually open question. Most daringly, he suggests that the way they coped with entanglements of science, politics and life is still representative of scientists now. By the end of this book, careful readers will be left with the queasy feeling that our own moral high ground has disappeared, and that Ball has revealed the ‘soul’ of physics to be no more intrinsically noble than any other. ■

Robert P. Crease is professor of philosophy at Stony Brook University, New York, and author of *World in the Balance*. e-mail: robert.crease@stonybrook.edu



BIOLOGY

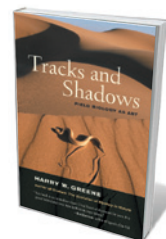
The love of pit vipers

Stuart Pimm follows a fellow biologist's evolution from wide-eyed wonder to a life chasing snakes in the field.

I shake the hand of my fellow guide on a tour group along the Amazon. Feeling missing fingers, I blurt out, “You’re a herpetologist?” Quickly forgiving me, he names the species of snake responsible. Harry Greene, in his engaging autobiography *Tracks and Shadows*, tells us of others who have lost digits. Greene himself still has a full set. He has been lucky — and careful.

We learn much about snakes from Greene, but more about the academic lineages and personalities that shaped his field. Greene and I are academic cousins, sharing a distant academic ancestor in the form of field biologist Joseph Grinnell, who worked at the University of California, Berkeley, from 1908 until his death in 1939. The theme of Greene's book is that the shadows cast by academic family mould our lives, but so do the species we track.

The field guide is the beginning. I vividly remember getting my first. It had to be of birds (my lifelong passion), because all Britain's amphibians and reptiles would form a



Tracks and Shadows: Field Biology as Art
HARRY W. GREENE
University of California Press: 2013.

small volume indeed. Only in graduate school in the American West, when my taxonomic passion was set, did I meet the groundbreaking guides to America's exceptional diversity of amphibians and reptiles by Robert Stebbins and Roger Conant. Greene flips through Conant's pages, and I imagine him thinking, “I want

to see that one. No, that's the one I just have to find!” And although the taxa differed, the experiences and outcomes were the same: we had to find what we saw that so intrigued us.

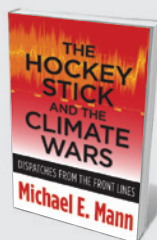
Two things follow. Soon, you are looking in places where you might find the real thing. Creeks, wetlands, woodlots and barren land — all places others might pass

MARTIN O'NEILL



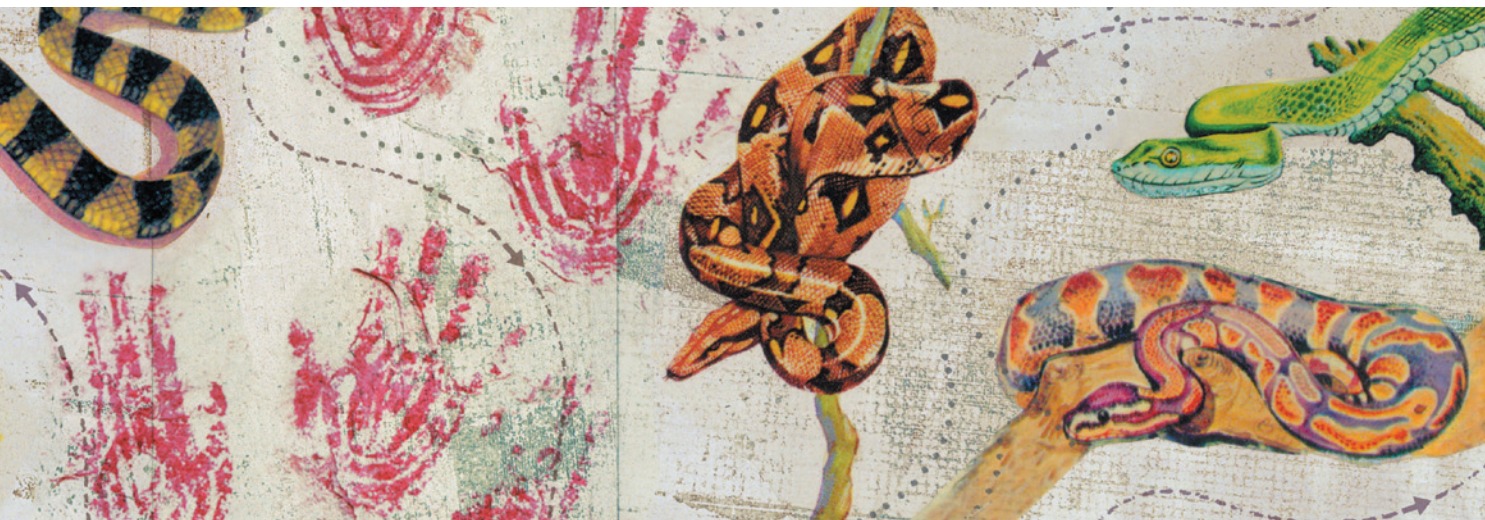
The War of the Sexes: How Conflict and Cooperation Have Shaped Men and Women from Prehistory to the Present

Paul Seabright (Princeton University Press, 2013)
An economist examines animals' tactics for ensuring reproduction, and ponders how human evolution can explain gender inequities in the West. (See John Whitfield's review: *Nature* **484**, 317; 2012.)



The Hockey Stick and the Climate Wars: Dispatches from the Front Lines

Michael E. Mann (Columbia University Press, 2013)
Meteorologist Michael Mann recounts the attack on his seminal 1998 global warming paper. The lengths to which deniers have gone to discredit the research continue to astound. (See Simon Lewis' review: *Nature* **483**, 402–403; 2012.)



without notice — become magical when one thinks they might hold desired species. Here, my path diverges from Greene's. I have never wished to turn over old planks to find rattlesnakes, to have my heart race as I pick up a cottonmouth, or to have to conceal being bitten by a copperhead from my parents.

The second discovery is that there are people just like you, with the same eccentricities, whose mentoring is vital. Greene writes a deeply respectful chapter about the herpetologist Henry Fitch, whom he met shortly after finishing high school. The sheer joy of learning more about natural history becomes an obsession, and mentors such as Fitch prove that it can be life-long. It carried Greene through an early job as a mortician's assistant, then the unavoidable Vietnam War years. As a medic, he looked after people battered and disabled by the conflict and contemplated his own imminent departure for battle. Sightings of whip-tailed lizards and black-tailed rattlesnakes near the training hospital provided welcome distractions. Seven of his colleagues were then posted to Vietnam and their fates still haunt him; he was posted to Germany.

After military service, graduate school was the University of Tennessee with a superb set of professors who, soon after Greene left, became my colleagues for 17 years. Greene's adviser, Gordon Burghardt, challenged him to think like a snake. "What are the *private*

experiences of animals?" Burghardt asks.

For the next two decades, Greene was at Berkeley, where he inherited Grinnell's desk. In those decades, technology made thinking like a snake easier. Radio transmitters revolutionized snake biology by allowing access to their secret lives.

Greene radio tracked rattlesnakes in deserts and bushmasters in rainforests, understanding what exceptional predators they are. We tend to view lions and tigers as

TRACKS AND SHADOWS IS AS PACKED WITH PEOPLE AND DRAMA AS A NOVEL.

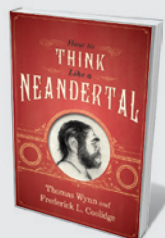
iconic hunters. Snakes, especially poisonous ones, are very different. They may sit and wait, catching prey only three to five times each year, yet must be ready to strike in a fraction of a second. Then, remarkably, they must use their tongue to sense the scent trail along which the fatally poisoned victim is fleeing. Digestion can take a week or more.

Sex is different too. Jesús Rivas, another of Burghardt's students, found his green

anacondas by feeling for them with his bare feet in the muck of the Venezuelan Llanos, a tropical grassland. Think like an anaconda: males are much smaller than the females. "Imagine lying for hours in ... a tropical slough, among a dozen seven-foot suitors for an eighteen-foot female, entangling your muscular, scaly tail with others competing for her vent." Males may need to be large enough to compete, but not so large as to be mistaken for a female, he explains.

Tracks and Shadows is as packed with people and drama as a novel, as Greene ventures forth with friends and revered mentors, records marriages and divorces, happiness and tragedies — some via snakebites — all uniquely wrapped in his herpetologist's world. As the 'art' in the subtitle indicates, he sees similarities between the immersive work of field biology and the worlds of the Amerindian rock artists of Texas and the painters of the caves at Chauvet in France — ancients who suffered life's vagaries in direct connection to the living world. Animals dominate as images. Modernity separates most of us from that life, but not so field biologists. ■

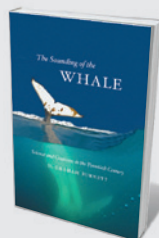
Stuart Pimm is professor of conservation at the Nicholas School of the Environment, Duke University, Durham, North Carolina, USA, and author of *The World According to Pimm: a Scientist Audits the Earth*. e-mail: stuartpimm@me.com



How to Think Like a Neandertal

Thomas Wynn and Frederick L. Coolidge (Oxford University Press, 2013)

This study of mental similarities between *Homo sapiens* and Neanderthals suggests that the powerful early humans had language, attended to their dead — and might have appreciated slapstick. (See Clive Gamble's review: *Nature* **479**, 294–295; 2011.)



The Sounding of the Whale: Science and Cetaceans in the Twentieth Century

D. Graham Burnett (University of Chicago Press, 2013)

Sobering insights abound in a history of cetacean science that powerfully reflects the mixed human response to Earth's largest mammal. (See Philip Hoare's review: *Nature* **481**, 141–142; 2012.)

ECONOMICS

Fixing the climate odds

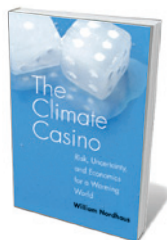
Gail Whiteman welcomes a take on climate economics that is strong on strategy.

The power of intelligent economics permeates William Nordhaus's *The Climate Casino*. In it, he presents an overview of climate science, economic theory and modelling, and outlines a number of economic strategies to resolve our climate challenges. He argues that economic growth is driving "unintended but perilous changes in the climate and earth systems" — and that we are, effectively, "rolling the climatic dice".

Not all may agree with this metaphor. But for US audiences in particular, the book convincingly makes the economic case for changing governmental policy, and our production and consumption habits, by offering economic incentives for low-carbon choices. The market alone cannot account for externalities stemming from climate change, such as ocean acidification, without being prodded by measures such as carbon taxes.

More debatable is what Nordhaus says about keeping the maximum global average temperature to 2°C above pre-industrial levels — a target of the Copenhagen Accord, the political compromise resulting from the 2009 United Nations Climate Change conference. Nordhaus views this goal as primarily political, and not well grounded in natural science, although numerous climate studies do support it. He suggests instead a rise of just under 3°C, as the review by Timothy M. Lenton and colleagues (T. M. Lenton *et al. Proc. Natl Acad. Sci. USA* **105**, 1786–1793; 2008) indicates that below it, large-scale tipping points such as widespread dieback in the Amazonian rainforest are unlikely.

But Nordhaus misses the point here. There is more to threshold setting than the avoidance of isolated tipping points. For example, the "planetary boundaries" model of Johan Rockström and others defines a "safe operating space" for humanity by pinpointing nine interlinked boundaries in Earth systems beyond which irreversible damage occurs (J. Rockström *et al. Ecol. Soc.* **14**, 32; 2009).



The Climate Casino: Risk, Uncertainty, and Economics for a Warming World
WILLIAM NORDHAUS
Yale University Press: 2013.

Although Nordhaus acknowledges this model's importance, he does not sufficiently integrate its range of critical boundaries into his own.

Nordhaus usefully differentiates between managed and unmanageable risks of climate impacts, underscoring the urgent need to prevent economies from triggering unmanageable risks from biodiversity loss,

for example. It is unfortunate, however, that the book's timing precludes the inclusion of reports by the Intergovernmental Panel on Climate Change Working Groups I and II (released last month and due in March 2014, respectively), as these would strengthen the chapters on climate science and impacts.

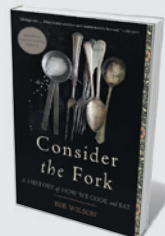
In his discussions on the strategies, costs, policies and institutions involved in slowing climate change, Nordhaus relies on his own integrated economic and geophysical model of climate-change economics, DICE (Dynamic Integrated model of Climate and the Economy). He offers a convincing comparison of carbon tax and cap-and-trade options, concluding that both are equally useful. And he argues strongly for a carbon price of US\$25 per tonne, using a 4% discount rate to bring future costs back to present-day dollars. However, other analysts support a much higher price for carbon and a lower discount rate, such as the one used in Nicholas Stern's groundbreaking 2006 review *Economics of Climate Change*. Nordhaus concludes: "We should aim for a lower temperature target if it is inexpensive, but we might have to live with a higher target if costs are high or policies are ineffective." Although

he is sensitive to the normative judgements of others, he does not perceive any normative sentiment in his own work.

Nordhaus's impassioned review of the politics around government (in)action on climate change and climate scepticism is largely US-centric. A more detailed analysis of the dramatic drop in European carbon prices in April 2012, following a decision by the European Parliament, would be welcome, for instance. What is interesting for all of us is Nordhaus's emphatic re-confirmation that his research using DICE does not support the position of climate sceptics, despite its use by some of these camps to argue that because climate change has economic benefits, there is no need to curb it.

Nordhaus is right in saying that economic incentives facilitate and encourage low-carbon behaviour. But managing climate change demands more. Markets are influenced by regulations and changes in accounting and valuation techniques that determine new rules of the game. The question of how best to deal with the thorny issue of stranded assets — obsolete or overvalued assets such as non-viable coal plants — remains unanswered in this book. Another missed opportunity is a deeper engagement with management theory, which has empirically shown that corporate behaviour across various industry sectors is driven by values, biases, emotions, culture and hyper-competitiveness as well as the pursuit of profit. Without delving deeper into corporate boardrooms, we are left wondering where change will come from if governments, as the architects of global policy frameworks, remain deadlocked. ■

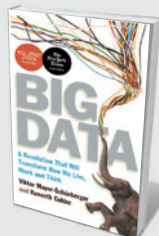
Gail Whiteman is professor-in-residence at the World Business Council for Sustainable Development in Geneva, Switzerland, and is Chair in Sustainability and Climate Change at Erasmus University, the Netherlands.
e-mail: gwhiteman@rsm.nl



Consider the Fork: A History of How We Cook and Eat

Bee Wilson (Basic Books, 2013)

Food historian Bee Wilson looks at how need sparks culinary innovation. She reveals, for instance, that China's lack of firewood led to the ultimate 'fast food' technique, stir-frying. (See Barbara Ketcham Wheaton's review: *Nature* **489**, 500; 2012.)



Big Data: A Revolution That Will Transform How We Live, Work and Think

Viktor Mayer-Schönberger and Kenneth Cukier (John Murray Publishers, 2013)

Big data is key to numerous fields and social-networking sites. Among many case studies, the authors contend that Google Flu Trends monitors influenza's spread better than traditional systems.



HISTORY OF SCIENCE

Science spun on the Silk Road

Christopher I. Beckwith assesses a study probing Central Asia's pivotal role in Islam's golden age.

Between Europe, the Near East, South Asia and East Asia lies a shockingly poor and underdeveloped region. But Central Asia — comprised mainly of Afghanistan, Uzbekistan, Turkmenistan, Tajikistan and East Turkistan (now Xinjiang) — was pivotal in pre-modern world history and cultural development, including science. Mathematician and astronomer al-Khwārizmī, for instance, systematized algebra, introduced decimal system mathematics and lent his name to algorithms (his Latinized name is Algorithmus). As Frederick Starr shows in *Lost Enlightenment*, Central Asia was a glittering, populous, wealthy world of advanced urban civilization in the mid-seventh century, when the first Arab armies reached Merv and Balkh, the “mother of cities”, in what are now,

respectively, Turkmenistan and Afghanistan.

Over the following decades, their armies crossed the Amu Darya (Oxus River) to Bukhara, Samarkand and Khwarizm. Less than two centuries later, the scholars of this region were mostly Muslim. They dominated the intellectual life of the entire Islamic world, stretching from Spain to India, and made fundamental contributions to the natural sciences, medicine, philosophy, music and literature. The philosopher al-Fārābī's *Great Book on Music*, for instance, became, as Starr writes, “the foundation stone of Western musicology”. And Western medicine was dominated until a few centuries ago by the works of al-Rāzī (Rhazes), the greatest clinical physician until early modern times, who was the first to precisely describe smallpox.

Starr argues rightly that the region's

brilliant culture rested on a highly cosmopolitan mix of ethnic groups, languages and religions; a long, rich pre-Islamic intellectual tradition (mainly Buddhist); and prosperity. That prosperity was built primarily on high-tech hydraulic engineering: Central Asians developed nine kinds of machinery for irrigation, drinking water and public baths. Soon after 1100 AD, the enlightenment waned under attacks on “reason and logic” led by the Sufi ex-philosopher al-Ghazālī.

At that point, medieval Western Europeans acquired science from the neighbouring Islamic world. They joined science to other Central Asian borrowings that institutionalized it and provided it with a formal scientific method that enabled it to survive and grow in Europe while science was dying in the Islamic world.

It is increasingly recognized that many of the greatest scientists, philosophers, poets and artists of the Islamic golden age were from Central Asia. A few of their works have been studied or translated, such as al-Birūnī's famous ethnography of India. But Starr's book is the first to identify the leading lights of that age as Central Asians, place them squarely in Central Asia, and detail their accomplishments.

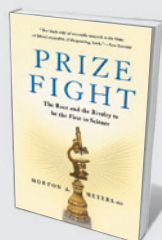
During the region's three centuries of world intellectual leadership, the dominant literary language was classical Arabic (except in East Turkistan, which became Islamic later). However, this was not due to the Arabs destroying Khwarizm's libraries, a claim repeated by Starr but shown by Wilhelm Barthold in 1928 to be folklore.



Lost Enlightenment: Central Asia's Golden Age from the Arab Conquest to Tamerlane

S. FREDERICK STARR
Princeton University Press: 2013.

In most of the world before the seventh century, people simply did not write much. Under the Arabs, the writing bug caught on and books in Arabic, and bookshops, became widespread in Central Asia. Starr relates how in the eleventh century, Ibn Sīnā (Avicenna) was chased down the street by a bookseller in Bukhara, eager to offer a bargain on an insightful ►



Prize Fight: The Race and the Rivalry to be the First in Science

Morton A. Meyers (Palgrave Macmillan, 2013)

A burning urge for discovery is often allied to a burning ambition for a Nobel. Among the cases here is that of Albert Schatz, who found streptomycin in 1943 but saw the prize go to his supervisor. (See Hidde Ploegh's review: *Nature* **486**, 318–319; 2012.)



Gravity's Engines

Caleb Scharf (*Scientific American*, 2013)

Astrobiologist Caleb Scharf investigates black holes — regions of space-time that pull in matter and light. He shows how those in galactic centres gobble stars, belch out plasma, and are the most efficient energy generators in the cosmos. (See Mario Livio's review: *Nature* **488**, 278; 2012.)

► volume about Aristotle's *Metaphysics* by the philosopher al-Fārābī. Ibn Sīnā later wrote many great works, including one of the most influential natural-science texts of the central Middle Ages, *De Visu* (*Optics*). This was translated into Latin in mid-twelfth century Toledo, Spain, by the Jewish philosopher Abraham ibn Daud and Dominicus Gundisalvi.

Linguistic unification by the Arabs meant that the flourishing of science and philosophy under Islam took place almost entirely in Arabic, as Starr suggests. Unfortunately, Starr uses his coinage “Persianate” throughout to refer specifically to the non-Persian peoples of Central Asia, making it sound as if the entire area was somehow “Persian” in language and culture. It was not. Persians, from what is now Iran, were conspicuously absent until the golden age was largely over, as Starr notes.

By calling his book *Lost Enlightenment*, Starr courageously rejects claims that there was no decline of Islamic civilization. He does, however, ignore recent work that explodes myths about Eurasian steppe peoples being aggressors, and even obliquely suggests that Chinggis Khan “attempted genocide” of Central Asians. Nevertheless, Starr firmly rejects the theory that the Mongols triggered the intellectual collapse. That, he writes, had happened a century before the Mongol conquest; at that time, taxes and trade were still “pouring gold into the coffers” of Central Asian rulers, who simply stopped using the money to support intellectual life. And after losing a great war — the Mongol ‘invasion’ (which historical sources agree the Khwarizmians started) — they failed to completely rebuild.

Starr shines in his core chapters, where he presents the great achievements of the Central Asian philosopher-scientists at a time when their homeland was the creative intellectual capital of the world. ■

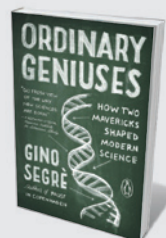
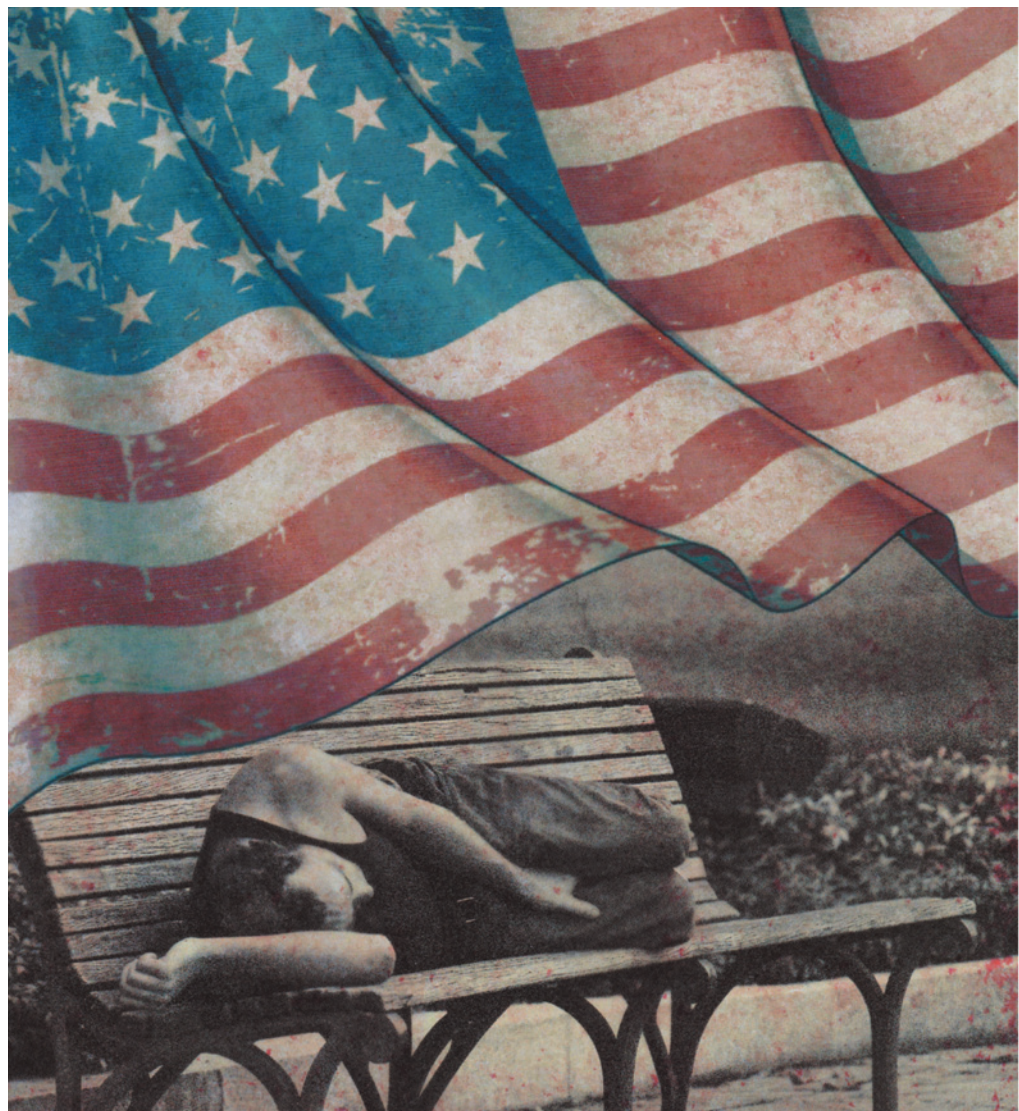
Christopher I. Beckwith is professor of Central Eurasian studies at Indiana University, Bloomington, and author of *Warriors of the Cloisters: The Central Asian Origins of Science in the Medieval World*.

e-mail: beckwith@indiana.edu

PSYCHIATRY

America the traumatized

Andrea Tone assesses a history of the mass release of US psychiatric patients into an uncertain future.



Ordinary Geniuses: How Two Mavericks Shaped Modern Science

Gino Segre (Penguin, 2013)

In these intertwined stories of cosmologist George Gamow and biologist Max Delbrück, we see how Gamow explained the creation of hydrogen and helium in the Big Bang, and Delbrück's study of bacterial viruses opened a new approach to genetics.



Memory: Fragments of a Modern History

Alison Winter (University of Chicago Press, 2013)

A subtly nuanced cultural and scientific history of our ‘recording mechanism’. Alison Winter reveals how memory has been tested variously in ‘labs’ like the courtroom, where phenomena such as false-memory syndrome have emerged. (See Barbara Kiser's review: *Nature* **479**, 475; 2011.)

In *American Psychosis*, E. Fuller Torrey turns to the past to determine why the United States has failed to care for the seriously mentally ill since de-institutionalization began in the mid-1950s. Between 1955 and 1969 alone, more than 220,000 patients were discharged from public psychiatric asylums. The scale of the problem this process has unfurled is visible today in parks, subway stations and emergency rooms where the under- and untreated go, partly because there are no other places for them. This serious issue deserves a one-two punch of compassion and political action.

Torrey, a psychiatrist, focuses on the federalization of mental-health care that began after the Second World War, when the National Institute of Mental Health was established (in 1946), federal grants were given to advance neuropsychiatric research, and outpatient community health centres were set up through the Community Mental Health Act of 1963. The number of beds in state-run institutions decreased as new medications, such as chlorpromazine (which contained the symptoms of illnesses such as schizophrenia), became available, and as families, politicians and activists sought to support patients outside the asylum. Torrey contends that this shift continues to fail psychiatric patients and that the state 'system' that federalization ostensibly usurped would have done better. He also argues that the Kennedy family, which has produced so many prominent US politicians, had a key role in this story.

Torrey begins his tendentious tale with Joseph P. Kennedy (1888–1969), businessman, social climber, diplomat and head of the clan. Torrey pinpoints what he regards as Kennedy's most serious failing: the decision to lobotomize his daughter Rosemary in 1941, after what was referred to as mild retardation became a major psychiatric disorder. According to Torrey, "mental retardation had been a family disgrace, but mental illness would be a debacle".

The result of that decision, Torrey argues, was a disaster for the future of the nation's mentally ill, not just for Rosemary. As he sees it, guilt over her lobotomy set the agenda of the family's political legacy, becoming "a family sin that demanded expiation". Decades later, in 1963, Joseph's son, President

John F. Kennedy, signed legislation to continue federalization with the establishment of publicly funded community mental-health centres. According to Torrey, by the end of 1976, 548 centres were running and almost 200 more had been funded. These, Kennedy stated in a speech to Congress, would spare the mentally ill the "cold mercy of custodial isolation" in state asylums. Torrey, however, avers that the centres were a flawed approach, based on the belief that serious psychiatric illnesses could be prevented or managed in outpatient centres.

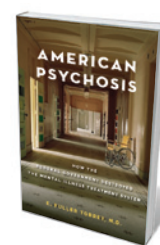
I disagree with much of this argument. First, the United States never had what Torrey refers to as a singular mental-illness treatment system. In the 1940s, it was just a maze of unevenly funded state public asylums. Their overcrowding, understaffing and often filthy conditions, and their cost to taxpayers drew public criticism and provided the impetus for political reform.

As historians have shown, lobotomies were a treatment of last resort, propelled by therapeutic nihilism, abominable conditions and the hope invested in new, radical, therapies. Fear that admission into a state institution might portend a life sentence of custodial care prompted families to authorize at least 20,000 lobotomies in the United

LOBOTOMIES WERE A TREATMENT OF LAST RESORT.

States between 1936 and the mid-1950s. To assert, as Torrey does, that if the federal government had not become involved, state hospitals would perhaps have provided something better, romanticizes what did not happen, while discounting the disturbing history that prompted federal intervention.

Also missing is a discussion of the influence of private hospitals on the demographics of psychiatric treatment since the mid-1950s. The affluent can access the best treatment; the poor are denied it. And by the 1960s and 1970s, as Jonathan Metzl's book, *The Protest Psychosis* (Beacon Press, 2010) shows, public



American Psychosis: How the Federal Government Destroyed the Mental Illness Treatment System
E. FULLER TORREY
Oxford University Press: 2013.

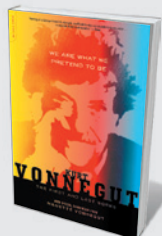
psychiatric hospitals were not a panacea, especially for African-American male patients, who came to represent a majority in many hospitals and whose treatment reflected racist views. Torrey also fails to discuss how the advent in the 1970s of private health plans provided by health maintenance organizations (HMOs) further impeded access to quality psychiatric care by offering financial incentives

to primary physicians to reduce referrals to specialists. HMOs and programmes to reduce health costs cemented a pattern that began in the 1950s. Now, increasingly, GPs make most front-line psychiatric diagnoses.

Torrey also ignores how a seismic shift in emerging psychiatric disorders, such as social phobia, has restructured psychiatric care. As less serious mental-health disorders such as mild depression became the therapeutic domain of psychiatry, such outpatient treatment claimed a larger part of psychiatrists' time, leaving less time and fewer institutions for patients battling serious illnesses with different needs.

In my opinion, *American Psychosis* fails to deliver a compelling explanation for the United States' present predicament, bogged down as it is in a tangle of initiatives — community, state and federal, public and private, medical and non-medical — and people in need. The book is nonetheless timely. It reminds us of the urgency of this problem and the need for fresh solutions to galvanize change. As Torrey contends, like President Kennedy before him, the nation's sick and most vulnerable citizens deserve better. ■

Andrea Tone is the Canada Research Chair in the Social History of Medicine in the Departments of History and Classical Studies, and Social Studies of Medicine at McGill University, and the author of *The Age of Anxiety*.
e-mail: andrea.tone@mcgill.ca



We Are What We Pretend To Be: The First and Last Works

Kurt Vonnegut (Vanguard Press, 2013)
The fiction of trained chemist Kurt Vonnegut touches on themes of societal ignorance and anti-authoritarianism. In this posthumous collection, Vonnegut's first and last pieces of fiction are pervaded by his trademark dark humour.



The Techno-Human Condition

Braden R. Allenby and Daniel Sarewitz (The MIT Press, 2013)

Technology is progressing so rapidly that we may be unable to fully prepare for it. This insightful take on a tangled issue points to the looming possibility of technological evolution outpacing human intent.

Ronald Harry Coase

(1910–2013)

Nobel-prizewinning economist whose work inspired cap-and-trade.

“I have had greatness thrust upon me,” Ronald Coase wrote to the committee that awarded him the Nobel Memorial Prize in Economic Sciences in 1991 at the ripe young age of 80 (he lived to be 102). This greatness was thrust upon Coase for his uncanny ability to think through important questions at the core of economics. At a University of Chicago conference in 2011 to celebrate his 100th birthday, the breadth of topics inspired by his work was remarkable: climate-change policy, the field of law and economics, economic development and telecommunications regulation.

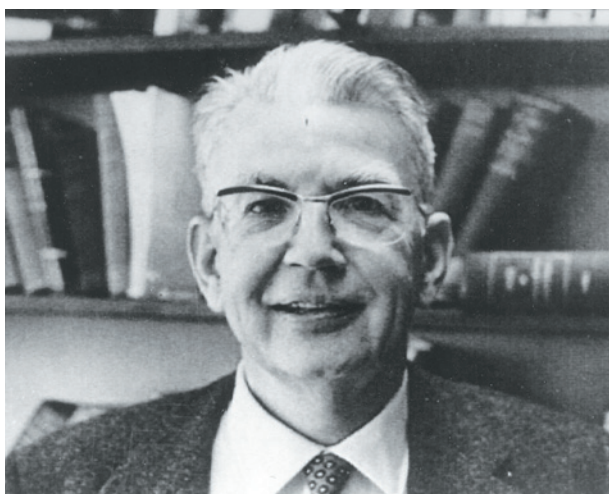
The Nobel committee rightly cited two extraordinary papers. To research the first, the British-born Coase left the London School of Economics in 1931 and spent a year in the United States. He visited factories and businesses to figure out why different industries were organized differently — such as barber shops with only a few employees and automobile companies with casts of thousands.

Aged 21, he returned to the United Kingdom and delivered a lecture in Dundee, arguing that companies exist because it is often cheaper to organize production that way. Having people on staff can save on transaction costs, such as having to repeatedly renegotiate labour contracts. Coase also relied on transaction costs to help explain why businesses do not grow forever — at some point they become too expensive to manage. Coase published his seminal paper on this subject, ‘The Nature of the Firm’ (R. H. Coase *Economica* 4, 386–405; 1937) six years later. He said he did not want to “rush into print” and had other teaching and research responsibilities — one of many examples of his humility.

In 1960 Coase published his masterpiece, one of the most cited, and arguably most misinterpreted, papers in economics: ‘The Problem of Social Cost’ (R. H. Coase *J. Law Econ.* 3, 1–44; 1960). It was based on arguments he had outlined the previous year in a paper on the US Federal Communications Commission (FCC), contending that the rights to use the electromagnetic spectrum should be bought and sold freely. This unfettered exchange would allow the spectrum to go to its most highly valued uses, which would be good for consumers.

Thus, mobile-phone networks could eventually displace television broadcasting as the demand for mobile-phone service increases. Today, Coase’s idea is conventional wisdom; at the time it was revolutionary.

Coase, then teaching at the University



of Virginia in Charlottesville, was invited to defend his FCC argument in front of a University of Chicago economics brain trust, including Milton Friedman and George Stigler. Coase won over his sceptical audience. In 1964, he accepted a professorship at the University of Chicago in Illinois, where he spent the rest of his career.

‘The Problem of Social Cost’ changed the way that economists think about externalities, such as pollution. Up to that point, it was generally believed that having government put a price on pollution, an idea advanced by the British economist A. C. Pigou 40 years earlier, was the best way to solve the problem. For example, a power plant might be asked to pay a US\$1 tax on each kilogram of sulphur dioxide it emits.

Coase argued for other possible solutions. He suggested that the overall level of harm from a factory is related to how close people choose to live to it, as well as to the smoke it emits. In this view, it is for both parties to minimize the overall damages from the pollution and the costs of avoiding those damages.

Coase suggested that polluters and their victims could achieve the socially efficient level of pollution through negotiations over who should pay for mitigation and what actions they should take — when two key

conditions hold. First, ownership of the property rights (in this case, to the environment) must be clearly defined; second, negotiations among parties must be costless. Under these conditions, and a few other technical assumptions, one gets the famous ‘Coase theorem’ (named as such by Stigler). This says that the initial distribution of property rights may not matter for achieving the socially efficient outcome. In 1990, policy-makers built on Coase’s insight in designing the cap-and-trade programme that cut US sulphur dioxide emissions by millions of tonnes.

Some analysts have taken the Coase theorem to suggest that government regulation is necessarily less efficient than private negotiations between parties, over the level of pollution, say. This is a misinterpretation. In some cases, negotiations will be better, typically when there are few affected parties so that negotiation costs are lower. In other cases, some kind of government intervention is likely to be more efficient, such

as in the regulation of greenhouse-gas emissions. Coase urged researchers to compare how different policy approaches might work in practice.

Coase believed strongly in understanding how institutions are built and sustained. He understood that markets — be they for derivatives, pork bellies or rights to emit carbon dioxide — do not come out of thin air. He encouraged his students, and their students, to learn about how markets form and why they work (or do not). As editor of the esteemed *Journal of Law and Economics* between 1964 and 1982, he encouraged careful empirical analyses of institutions and regulations. In 2000, he helped to launch the Coase Institute, based in St Louis, Missouri, which assists outstanding young scholars studying economic and political institutions.

Coase was a vocal critic of ‘blackboard economics’, in which equations are used to model economies that bear little resemblance to real-world organizations. Today his view is heretical in many mainstream economics departments. We ignore it at our peril. ■

Robert Hahn is director of economics and a professor at the Smith School of Enterprise and the Environment, University of Oxford, UK.
e-mail: robert.hahn@smithschool.ox.ac.uk

Small-brained and big-mouthed

A complete hominin cranium found at the archaeological site of Dmanisi shows remarkably primitive morphology, prompting its discoverers to propose that early forms of the genus *Homo* evolved as a single, highly variable lineage.

FRED SPOOR

Hominins are species more closely related to humans than to chimpanzees. The oldest hominin fossils that have been found outside Africa are from 1.77-million-year-old strata at Dmanisi in the Georgian Caucasus¹. These specimens show closest similarities to early *Homo erectus* fossils², but the discovery in 2000 of a large and robust mandible at the site led to the Dmanisi hominins being attributed to a new species, *Homo georgicus*³, with this mandible as its holotype. But other researchers have considered whether the size and shape differences between Dmanisi mandibles indicate that these fossils represent more than one species^{4,5}. Writing in *Science*, Lordkipanidze *et al.*⁶ now provide the description and comparative analysis of the cranium associated with the large *H. georgicus* mandible — and use this to infer a taxonomy that breaks with a decades-long consensus on hominin evolution.

The new specimen is complete, and together with the mandible it forms the best-preserved adult skull of a hominin from the Early Pleistocene (the period from around 2.6 million to 0.8 million years ago). This exceptional find is made even more exciting by the fact that limb bones have been recovered that probably belonged to the same individual, and that it can be considered in the context of four other crania that were previously uncovered from the same location and that belong to broadly the same time period.

The latest cranium is characterized by a large and projecting face, combined with a brain size of 546 cubic centimetres, which is smaller than those of the other Dmanisi crania and any other specimen attributed to *H. erectus* (Fig. 1). In their analyses, Lordkipanidze *et al.* consider two broad issues: whether the variation shown by the five Dmanisi crania is greater than expected for a single species, and what the implications of this cranial variation are for the interpretation of the early fossil record of the genus *Homo*. With respect to the first question, the authors find the overall cranial shape variation of the Dmanisi sample to be consistent with that seen in chimpanzees or modern humans, such that it can be accommodated within a single species.

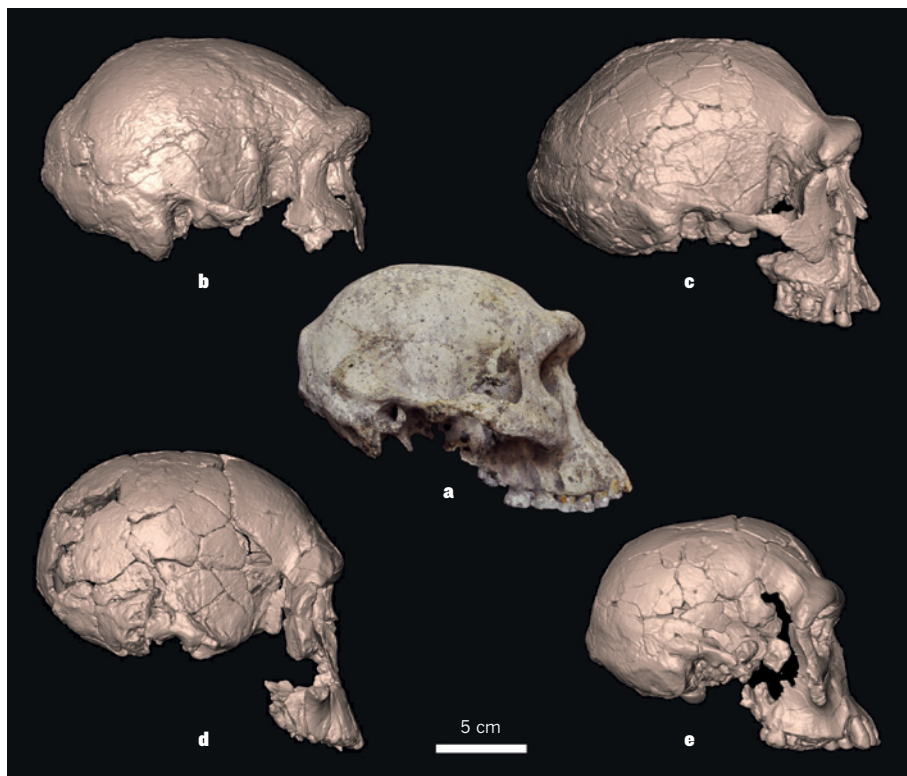


Figure 1 | A new cranium from Dmanisi. Lordkipanidze *et al.*⁶ describe a fossil cranium from the Dmanisi site in Georgia, known by its accession number, D4500. Here, it is presented in comparison with early *Homo* specimens from Kenya, dated to between 1.6 million and 2 million years ago. The new cranium (a) has a projecting face and a braincase that is small but similar in shape to those of *Homo erectus* specimens KNM-ER 3833 (b) and KNM-ER 3733 (c). Specimens KNM-ER 1470 (d) and KNM-ER 1813 (e), which are attributed to the species *Homo rudolfensis* and *Homo habilis*, respectively, differ from these three crania in the shape of the braincase or face. Lordkipanidze and colleagues argue that the differences between all early *Homo* specimens, including the five shown here, can be accommodated within a single species, *H. erectus*. Surface reconstructions (b–e) derived from computed tomography images (c, left side, reversed).

The second issue focuses on whether the diversity of early *Homo* fossils reflects an evolutionary radiation of multiple species (*Homo habilis*, *Homo rudolfensis* and *Homo erectus*)^{7–9}, or a single, highly variable lineage¹⁰. On the basis of cranial shape analyses and a broad comparison of characteristics, the authors report that the morphological variation seen in the African fossil record of early *Homo* lies within the variation shown by chimpanzees, modern humans or the Dmanisi sample. This leads them to conclude that early *Homo* evolved as a single variable lineage, and to attribute the associated fossil record to a single

species, *H. erectus* (this name has priority over others because it was the first one used for any of these fossils). Consequently, the authors retract *georgicus* as a species name, but re-use it in their designation of the Dmanisi sample as *Homo erectus ergaster georgicus*. This highly unusual infrasubspecific classification is probably the first use of a quadrinomial in primate taxonomy, and is not recognized by the International Code of Zoological Nomenclature.

The radical proposal to subsume the well-established taxa *H. habilis* and *H. rudolfensis* into *H. erectus* warrants careful scrutiny, and in my view the presented evidence is weak. It

is doubtful whether analyses of overall cranial shape have the diagnostic power to distinguish between closely related taxa, as is indeed demonstrated by some of the analyses presented in the report. Species are defined by specific morphological features, not by overall cranial shape. Lordkipanidze and colleagues' list of individual features could have been informative in this respect, but it is not analysed systematically, nor is a distinction made between traits that are derived (absent in the last common ancestor of a group) or primitive (already present in the last common ancestor) — a distinction that is essential to establishing phylogenetic relationships. Moreover, the features are categorized in a way that sometimes obscures, rather than highlights, important differences. For example, two crania attributed to *H. rudolfensis*⁹ clearly differ from all other early *Homo* specimens in the degree of facial projection around the mouth. This distinction is not revealed in the authors' table of features because of the arbitrary way the associated angle is categorized. Finally, the authors make no reference to the available non-cranial fossil evidence, even though biomechanical analyses of specimens attributed to *H. habilis* and *H. erectus* indicate marked differences in locomotive behaviour¹¹.

The new cranium's small brain size, projecting face and large cheek teeth are primitive for *H. erectus* (in the conventional use of this species name), but the specimen also shows derived morphological features that are typically found in this species, but not in specimens attributed to *H. habilis* or *H. rudolfensis*. These include its thick and protruding brow ridges, the distinct shape of the occipital bone (Fig. 1) and the arrangement of the temporal bone in basal view. This pattern of combined primitive and derived morphology is seen in other Dmanisi specimens as well, but in the new cranium the primitive aspect is particularly prominent. As such, this morphology seems to correspond to what one would expect not too long after the *H. erectus* lineage diverged from a more generalized form of early *Homo*. It would also be compatible with the centrifugal model of speciation¹², in which central populations in Africa are more derived, and peripherally distributed ones in western Asia and southern Africa (such as *Homo* at the Swartkrans site) retain primitive features.

The discovery of the new Dmanisi cranium will greatly help with the evaluation of the fossil record of early *Homo* in eastern Africa, which is temporally and geographically more

diverse, and generally less well-preserved. This should contribute to a better understanding of where and when the *H. erectus* lineage first emerged, and how it relates to other taxa of early *Homo*. ■

Fred Spoor is in the Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany, and University College London, UK. e-mail: f.spoor@eva.mpg.de

1. Lordkipanidze, D. et al. *Nature* **449**, 305–310 (2007).
2. Rightmire, G. P. & Lordkipanidze, D. in *The First Humans* (eds Grine, F. E., Fleagle, J. G. & Leakey, R. E.) 39–48 (Springer, 2009).
3. Gabunia, L., de Lumley, M. A., Vekua, A., Lordkipanidze, D. & de Lumley, H. C.R. *Palevol* **1**, 243–253 (2002).
4. Skinner, M. M., Gordon, A. D. & Collard, N. J. *J. Hum. Evol.* **51**, 36–49 (2006).
5. Rightmire, G. P., Van Arsdale, A. P. & Lordkipanidze, D. *J. Hum. Evol.* **54**, 904–908 (2008).
6. Lordkipanidze, D. et al. *Science* **342**, 326–331 (2013).
7. Wood, B. *Nature* **355**, 783–790 (1992).
8. Spoor, F. et al. *Nature* **448**, 688–691 (2007).
9. Leakey, M. G. et al. *Nature* **488**, 201–204 (2012).
10. Suwa, G. et al. *Anthropol. Sci.* **115**, 133 (2007).
11. Ruff, C. Am. J. Phys. *Anthropol.* **138**, 90–100 (2009).
12. Groves, C. P. *A Theory of Human and Primate Evolution* (Clarendon, 1989).

ASTROPHYSICS

Recipe for regularity

A detailed astrophysical model has been laid out that not only reproduces the far-infrared-radio correlation for galaxies that are actively forming stars, but also predicts how the correlation is modified at high redshift.

ELLEN ZWEIBEL

Galaxies, particularly those that, like our own Milky Way, actively form stars, are complex systems in which a vast array of physical processes operate simultaneously (Fig. 1). Patterns of regularity in galaxy behaviour are often interpreted, therefore, as evidence for global self-organization, and thus for the workings of the system at a fundamental level. One such pattern is a remarkably tight correlation between the rate of star formation and that of synchrotron radiation from cosmic-ray electrons gyrating in the galactic magnetic field. This correlation has now been reinterpreted by Schleicher and Beck, through the lens of modern ideas about magnetic-field amplification in galaxies, in a paper¹ published in *Astronomy & Astrophysics*. The authors provide predictions about the evolution of the correlation, and the physical quantities underlying it, over cosmic time.



Figure 1 | An ultra-luminous infrared galaxy. Galaxy IRAS 19297-0406, shown here in a composite image, is an extreme example, in terms of its star-formation rate, of the type of galaxy described by Schleicher and Beck¹.

These predictions will be testable with the radio telescopes currently under development.

Observations at wavelengths from radio to γ -rays are providing a detailed picture of the structure and evolution of galaxies, from their formation just a few hundred million years after the Big Bang to the present. One of the most tantalizing results to emerge from multi-wavelength studies is that the far-infrared (FIR) and radio luminosities (L_{fir} and L_{rad}) from galaxies over several orders of magnitude in galaxy luminosity and size show a weakly nonlinear correlation: $L_{\text{rad}} \sim L_{\text{fir}}^x$, where the exponent x has a value in the range 1.15–1.3 (refs 2,3).

The radio luminosity is primarily emitted by relativistic cosmic-ray electrons circling galactic magnetic-field lines and is roughly proportional to the product of magnetic-field and cosmic-ray-electron energy densities. The FIR luminosity is emitted by interstellar dust heated by the ultraviolet radiation from massive stars. Because the lifetimes of these stars are short by galaxy-evolution standards (a few million years), the number of these stars in a galaxy is proportional to the rate at which they form. Thus, the FIR–radio correlation suggests that the product of magnetic-field and cosmic-ray-electron energy densities scales with the star-formation rate, with exponent x and a scatter of only about 2 over a wide range of galaxy properties. The sensitivity and resolution of telescopes have now improved to the point

that this correlation has been confirmed in galaxies that have cosmological redshifts of about 2, which we observe at a time when the Universe was only about one-fifth of its present age.

A model illustrates the plausibility of the FIR–radio correlation. Because massive stars end their lives as supernovae, the supernova rate scales with the star-formation rate. There is good evidence that cosmic rays are accelerated by supernovae. Suppose a fraction of the energy of each supernova is converted to cosmic rays. Suppose further that the main energy sink for cosmic rays is synchrotron radiation. Then, the energy density of cosmic-ray electrons is directly proportional to the star-formation rate and inversely proportional to the magnetic-energy density, whereas the synchrotron emissivity is independent of magnetic-energy density and directly proportional to the star-formation rate. So, by assumption, is the FIR emissivity; therefore, the synchrotron and FIR emissivities are correlated.

This model is a simplified version of so-called calorimeter models of the FIR–radio correlation⁴. More general versions, which include mechanisms of electron-energy loss other than synchrotron radiation, such as inverse Compton scattering of electrons by ambient photons or electron escape from the galaxy, do yield synchrotron emissivity that depends on magnetic-energy density. This introduces an element of uncertainty into the models, because galactic magnetic fields are difficult to measure, and the theory of how

they originated and grow is still incomplete⁵.

In their paper, Schleicher and Beck have laid out a more detailed model that reproduces the observed FIR–radio correlation. The model's new ingredients are an estimate of galactic magnetic-field strength based on recent results⁶ from the theory of magnetic-field amplification by galactic turbulence, and an estimate of the level of galactic turbulence which ties it to the star-formation — or supernova — rate. There are good theoretical and empirical bases for both estimates. Although the origin of the large-scale magnetic fields seen in many galaxies is still unclear, the idea that turbulence regulates the amplitude of a small-scale turbulent magnetic field such that the energy densities of the two are proportional is well established⁶. What matters for the FIR–radio correlation is magnetic-energy density, not large-scale field structure. Likewise, it has long been argued, on general energetic grounds, that energy supplied by massive stars and supernovae is a primary driver of turbulence in the interstellar medium. But up to now, these well-founded ideas had not been used quantitatively in a model of the FIR–radio correlation.

On the basis of their simple models of turbulence driving and magnetic-field amplification, Schleicher and Beck derive a weak nonlinearity of the FIR–radio correlation. Galaxies with low star-formation rates have less turbulence, weaker magnetic fields, smaller synchrotron losses and lower radio fluxes; at high star-formation rates the opposite holds. The model also

predicts how the FIR–radio correlation should scale with cosmological redshift. The density of cosmic microwave background photons that permeate the Universe and mean star-formation rates increase with redshift, enhancing the importance of inverse Compton emission relative to synchrotron emission at high redshift (inverse Compton emission is produced by cosmic-ray electrons interacting with photons, so if there are more cosmic-background and starlight photons there is more inverse Compton emission). This alters the correlation at high redshift, a change that should be observable with the Square Kilometre Array (SKA) radio telescope now under construction. Verification of Schleicher and Beck's prediction would be evidence for rapid turbulent amplification of magnetic fields in the early lives of galaxies. ■

Ellen Zweibel is in the Departments of Astronomy and Physics, University of Wisconsin, Madison, Wisconsin 53706, USA. e-mail: zweibel@astro.wisc.edu

1. Schleicher, D. R. G. & Beck, R. *Astron. Astrophys.* **556**, A142 (2013).
2. de Jong, T., Klein, U., Wielebinski, R. & Wunderlich, E. *Astron. Astrophys.* **147**, L6–L9 (1985).
3. Helou, G., Soifer, B. T. & Rowan-Robinson, M. *Astrophys. J. Lett.* **298**, 7–11 (1985).
4. Völk, H. J. *Astron. Astrophys.* **218**, 67–70 (1989).
5. Kulsrud, R. M. & Zweibel, E. G. *Rep. Prog. Phys.* **71**, 046901 (2008).
6. Tobias, S. M., Cattaneo, F. & Boldyrev, S. in *Ten Chapters in Turbulence* (eds Davidson, P. A., Kaneda, Y. & Sreenivasan, K. R.) 351–404 (Cambridge Univ. Press, 2013).

could be synthesizing a signalling molecule that, when secreted, promotes fat uptake by the muscle. Indeed, they find that blood serum collected from normal mice in the dark phase of the day can promote fat uptake by cultured muscle cells, but that serum from mice lacking PPAR δ in the liver cannot.

Extensive analysis narrowed down the factors transmitting the effects of PPAR δ through the blood to a handful of lipid candidates, and Liu *et al.* focused on a phosphatidylcholine dubbed PC(18:0/18:1), demonstrating that treatment with this phospholipid, but not with other closely related phosphatidylcholine species, induces fatty-acid uptake into muscle cells both *in vitro* and *in vivo*. This is a hallmark of PPAR α activation, and, consistently, PC(18:0/18:1)-mediated fatty-acid uptake was diminished in PPAR α -deficient muscle cells and in mice.

Thus, this dance starts at night when liver PPAR δ is activated, increasing PC(18:0/18:1) production. In an exchange of partners, PC(18:0/18:1) crosses from the liver to muscle, where it joins with PPAR α in the next step, promoting fat uptake and fatty-acid oxidation. The cycle is completed as the levels or activities of all three partners fall during the day, setting up the next round.

Now that they have been worked out, these

PHYSIOLOGY

A metabolic minuet

Two related nuclear receptors mediate circadian fat metabolism in two different tissues using a lipid messenger as an intermediary. This signalling pathway might be relevant to the understanding of metabolic disorders. SEE LETTER P.550

DAVID D. MOORE

In the minuet, a popular court dance of the baroque era, couples exchange partners in recurring patterns. This elaborately choreographed exercise comes to mind when reading Liu and colleagues' paper¹ on page 550 of this issue. In this study, the nuclear receptors PPAR α and PPAR δ are two of the three stars in a metabolic minuet that promotes appropriate fat utilization.

PPAR α drives fat use in muscle and liver and is a well-known target of the fibrate class of lipid-lowering drugs. By contrast, PPAR γ is essential for the development of white-fat tissue, mediating fat storage. PPAR δ is more broadly expressed than its two brothers and is more enigmatic, having functions that overlap with both. In muscle it promotes fatty-acid

breakdown and increases muscle endurance^{2,3}. And in the liver, it stimulates fatty-acid synthesis, or lipogenesis, as Liu and co-workers have previously demonstrated⁴. This lipogenic activity is now shown to generate a 'dancing partner' for PPAR α .

The circular pattern for this dance comes from the circadian activity of PPAR δ in the liver (Fig. 1). Mice eat at night, storing excess calories as fat. During the day, Rev-erba and Rev-erbb, two nuclear receptors that also have circadian activity, repress lipogenesis in this organ⁵. Liu *et al.* report that nocturnal expression of at least a subset of key lipogenic enzymes in the liver depends on PPAR δ . They also make the surprising observation that mice lacking PPAR δ in the liver have defective fat uptake in muscle, but only at night. The authors deduce that the night-time liver

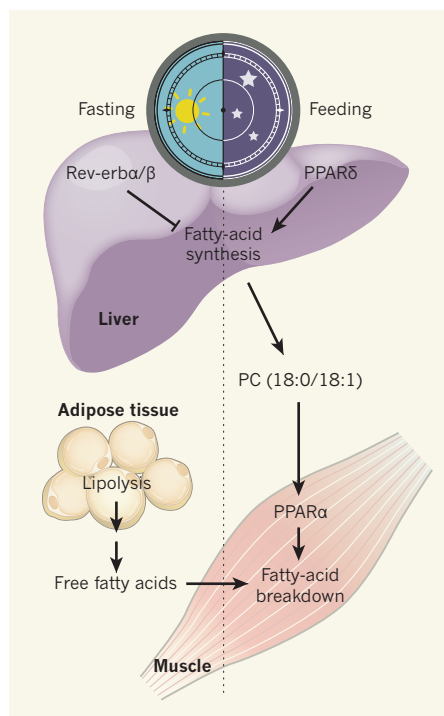


Figure 1 | Cross-tissue regulation of fat metabolism. Mice store excess calories from their nocturnal feeding as fat, and synthesize fatty acids in the liver. The nuclear receptors Rev-erba/β repress this process during the day. Liu *et al.*¹ show that PPARδ promotes night-time lipogenesis in the liver. The phospholipid PC(18:0/18:1) then moves to peripheral tissues such as muscle, where the related nuclear receptor PPARα mediates fatty-acid breakdown. Lipolysis in adipose tissue fuels the muscle.

dance-like steps might seem relatively simple. Yet their potential importance is highlighted by the authors' observations that circadian production of PC(18:0/18:1) is dampened in mice fed a high-fat diet, and that PC(18:0/18:1) treatment improves metabolic parameters in diabetic mice, modestly decreasing blood levels of triglycerides and improving glucose homeostasis. Overall, these results are consistent with the beneficial effects of PPARα-activating fibrate drugs. They also suggest that the time of day at which fibrate treatment is given might be important, and that a drug that specifically targets PPARδ could still have PPARα-mediated side effects.

The new data also raise a host of difficult but intriguing questions. For instance, why does fatty-acid production in the liver promote the opposite process of fatty-acid oxidation in skeletal muscle? A more tractable question is whether PC(18:0/18:1) directly activates muscle PPARα. The answer is probably yes, given previous observations⁶ that other phosphatidylcholines can also activate PPARα and that the nearly identical PC(16:0/18:1) is a highly specific ligand for PPARα in the liver⁷. However, Liu *et al.* report that PC(16:0/18:1) does not activate PPARα in muscle cells. The reason for this apparent discrepancy is not obvious,

and the nature of the endogenous functional activators of all three PPARs remains unclear. Extensive functional, biochemical and structural studies are needed to fully address this long-standing question.

Both PC(18:0/18:1) and PC(16:0/18:1) are abundant components of cell membranes. This raises the broader question of how such common molecules could function as specific metabolic signals. It could be that cellular compartmentalization is involved, such that the phospholipids that signal in the nucleus are somehow separated from the same molecular species in the cell membrane.

Several studies from another lab^{7–9} have suggested a specific compartmentalization pathway in which the enzyme fatty-acid synthase is required for production of the endogenous PPARα ligand in the liver. In response to nutrient signals, this pathway channels lipid synthesis through specific subcellular compartments to generate nuclear PC(16:0/18:1); only newly minted phosphatidylcholine is active in this scenario. The lipogenic component of the PC(18:0/18:1) story provides an intriguing parallel. Unfortunately, however, the idea that only newly produced intracellular phosphatidylcholine is active is not consistent with the biological effects of exogenously added PC(16:0/18:1) described previously⁷, nor with those of PC(18:0/18:1) in the current study. It is not clear how PC(18:0/18:1) exerts its effects in skeletal muscle, nor how it avoids PPARα activation in the liver, which would counteract the effects of PPARδ in a futile cycle of coincident fat synthesis and oxidation.

APPLIED PHYSICS

Materials scientists take control

The discovery of a new way of controlling a class of complex-oxide materials, known as the Ruddlesden–Popper series of structures, may lead the way to making electronically tunable microwave devices. SEE LETTER P.532

MELANIE W. COLE

One might say that materials scientists are control freaks, because they are constantly seeking new strategies to control and improve material properties, and to manipulate materials to create new functionalities. This is especially true for the field of thin-film complex-oxide materials and their development for microwave electronics. So far, barium strontium titanate (BST) thin films look like the most promising complex-oxide material for developing inexpensive, small, electronically tunable microwave

And a final question concerns generality. If this PPAR dance is the minuet, what about the gavottes and rigadoons, to say nothing of the square dances? The intracellular regulatory effects of lipid signalling molecules such as diacylglycerol and ceramides are well known, and release of the specific lipid-controlling hormone C16:1n7-palmitoleate from adipose tissue promotes insulin action in muscle and suppresses fat accumulation in the liver¹⁰. More in line with the PPARδ–PC(18:0/18:1)–PPARα interchange, the nuclear receptors SF-1 and LRH-1 respond to phospholipid ligands^{11–13} to exert direct metabolic effects. Clearly, we don't know all the steps that the dance master has choreographed. ■

David D. Moore is in the Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA. e-mail: moore@bcm.edu

1. Liu, S. *et al.* *Nature* **502**, 550–553 (2013).
2. Wang, Y.-X. *et al.* *Cell* **113**, 159–170 (2003).
3. Narkar, V. A. *et al.* *Cell* **134**, 405–415 (2008).
4. Liu, S. *et al.* *J. Biol. Chem.* **286**, 1237–1247 (2011).
5. Feng, D. *et al.* *Science* **331**, 1315–1319 (2011).
6. Lee, H. *et al.* *Circ. Res.* **87**, 516–521 (2000).
7. Chakravarthy, M. V. *et al.* *Cell* **138**, 476–488 (2009).
8. Chakravarthy, M. V. *et al.* *Cell Metab.* **1**, 309–322 (2005).
9. Jensen-Ustad, A. P. L. *et al.* *J. Lipid Res.* **54**, 1848–1859 (2013).
10. Cao, H. *et al.* *Cell* **134**, 933–944 (2008).
11. Urs, A. N., Dammer, E. & Sewer, M. B. *Endocrinology* **147**, 5249–5258 (2006).
12. Lee, J. M. *et al.* *Nature* **474**, 506–510 (2011).
13. Blind, R. D., Suzawa, M. & Ingraham, H. A. *Sci. Signal.* **5**, ra44 (2012).

devices that have high performance and low power consumption¹. But practical tunable-device applications demand maximum film tunability together with minimum dielectric loss (minimum signal attenuation)². Unfortunately, Mother Nature is not always amenable to our wants: for BST, she has configured these properties such that loss and tunability are negatively opposed to one another. On page 532 of this issue, Lee *et al.*³ describe an approach* to achieving an improved balance of these two properties which may lead to

*This article and the paper under discussion³ were published online on 16 October 2013.



50 Years Ago

Much excellent archaeological work has been done on Stonehenge ... It has been established that there was building activity from approximately 2,000 B.C. until 1,500 B.C. At the beginning of this period the 56 Aubrey holes were dug at equal spacings around a circle with errors of less than 0.5°. At the final phase the giant trilithon archways were in position, surrounded by the sarsen circle ... Positions of all stones, holes and midpoints were measured ... The machine programme called for the positions of stones, stone holes etc., in selected pairs, and the azimuths and horizontal declinations were computed. These alignments were then compared with the positions of the celestial bodies, and the errors of alignment computed. Stars and planets yielded no detectable correlation ... The Sun yielded 10 correlations; to a mean accuracy of 1.5° the Moon gave 14.

From *Nature* 26 October 1963

100 Years Ago

The Gypsy Lore Journal is largely devoted to an account by Mr. E. O. Windstedt of "The Gypsy Coppersmiths' Invasion of 1911–1913." Owing to the reticence displayed by these people, the origin of the party which visited England is uncertain. ... They appear to be genuine Gypsies, their skin colour being practically identical with that of the Russian peasantry. In their metal work there are remarkable coincidences with Indian art products. This monograph contains a very complete account of their religious beliefs, organisation, dress, manners, and customs. The excellent work being carried out, with very limited resources, by the Gypsy Lore Society ... should invite support from all who are interested in this remarkable race and from students of anthropology.

From *Nature* 23 October 1913

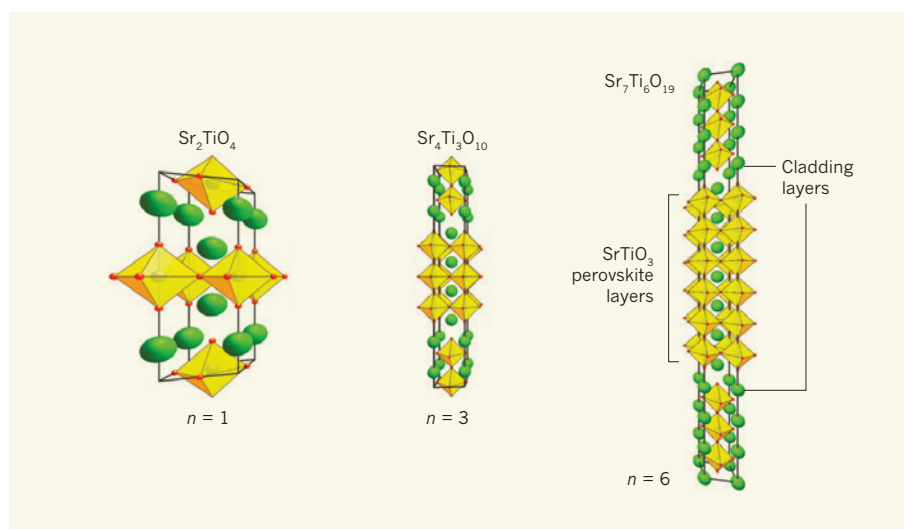


Figure 1 | The Ruddlesden–Popper series of structures. The $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ crystal structures studied by Lee *et al.*³ consist of perovskite SrTiO_3 layers sandwiched between SrO cladding layers. Structures with $n = 1, 3$ and 6 are displayed. Strontium atoms are represented as green spheres; titanium atoms are in the centre of the octahedra (yellow), with oxygen atoms (red spheres) at each apex.

the ultimate material for tunable microwave devices.

The authors' approach was to begin with an inherently low-loss dielectric (insulator) material system that is related to BST, namely, the Ruddlesden–Popper series of structures, $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$, and to engineer it to improve its tunability. To appreciate the engineered design that the authors have achieved at the atomic level, one needs first to visualize the structure of these materials. They are composed of perovskite SrTiO_3 layers situated between terminal SrO cladding layers (Fig. 1). These structures have been known for more than 50 years, but they are essentially dead when it comes to electronic tunability. Two years ago, theorists predicted⁴ that, under biaxial tensile strain, a ferroelectric structural instability — which consists of the cooperative motion of each positively charged titanium cation moving against the surrounding negatively charged octahedron of oxygen anions — emerges in the Ruddlesden–Popper phases. Such ferroelectric instability is exactly what is responsible for the electronic tunability in BST.

The unusual prediction⁴ about these Ruddlesden–Popper phases, however, is that this ferroelectric instability is local to each SrTiO_3 layer and occurs only if the spacing or distance between the terminal SrO cladding layers is large enough. In other words, the insertion of a specified number, n , of SrTiO_3 layers will increase the distance between the two terminal cladding SrO layers; and at some critical value of n , a ferroelectric instability will occur and with it electronically tunable behaviour attained through the application of an electric field.

Lee and colleagues' theoretical calculations show that for n greater than 3 (that is, for three or more perovskite SrTiO_3 layers inserted in between the two SrO layers), a

local ferroelectric instability occurs for these Ruddlesden–Popper films in which the crystal lattice is strained to match that of the underlying dysprosium scandate (DyScO_3) substrate. This innovation is particularly exciting to the materials-science community because there is now, for the first time, a control parameter, n , that can be used to manipulate the properties of a tunable dielectric to satisfy the low-loss and high-tunability demands required for electronically tunable microwave devices such as filters, delay lines and phase shifters. What's more, the approach does not involve adding undesirable atomic disorder to the system, which would increase its dielectric loss.

Armed with this theory, Lee *et al.* set out to test it experimentally. They not only validated the theory but also demonstrated that the temperature (T_c) at which the material undergoes a structural phase change, from the 'paraelectric' state above T_c to one with local ferroelectric order below T_c , could be manipulated by changing n , and that the SrO cladding layers serve to accommodate film non-stoichiometry. The latter discovery is particularly important because non-stoichiometric behaviour is usually accommodated by undesirable structural point defects that unfavourably enhance the material's dielectric loss. This alternative accommodation of film non-stoichiometry preserves the films' low-loss attribute. Detailed experiments revealed that $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ with $n = 6$ exhibited low loss and good tunability that was stable over a broad operational frequency range (1 kilohertz to 125 gigahertz). This behaviour is of paramount importance, because it shows that tunable devices composed of these films are frequency agile — that is, they can be used over a wide spectrum of frequencies with stable, predictable and enhanced performance.

As with all discoveries, there will always be naysayers who will not recognize new findings

until their intrinsic worth trumps the current best-in-class technology — in this case, BST. When comparing negatively opposed properties, such as loss and tunability, between different material systems, it is useful to have a figure-of-merit (FOM), a quantity used to characterize the performance of the material system. The most widely accepted FOM for tunable dielectrics is the ratio of tunability to dielectric loss. This FOM reflects the fact that a tunable microwave circuit cannot take full advantage of high tunability if the loss is too high⁵. Lee *et al.* present beautiful data showing that the FOM over the frequency range of 1–125 GHz for these new films is significantly better than that of BST.

So where do we go from here? For starters, the discovery of this new control parameter should motivate research to see if it can be used in related systems, to achieve even higher performance. The transition of this material into practical devices will also require replacement of the expensive, microwave-friendly, research-scale (small-size) designer DyScO₃ substrate with a large-area, low-cost, microwave-friendly substrate. However, Lee and colleagues have offered a jump-start for this by suggesting a relaxed (non-strained) DyScO₃ buffer layer on microwave-relevant substrates as a viable solution. Regardless of the path forward, this discovery, which allows a ferroelectric instability to be tuned by atomic

engineering without adding disorder, is exciting and unlocks possibilities for using atomic engineering to override Mother Nature's lack of materials-science practicality. ■

Melanie W. Cole is at the US Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005, USA
e-mail: melanie.w.cole.civ@mail.mil

1. Cole, M. W. *et al.* *Appl. Phys. Lett.* **92**, 182906 (2008).
2. Bao, P., Jackson, T. J., Wang, X. & Lancaster, M. J. *J. Phys. D* **41**, 063001 (2001).
3. Lee, C.-H. *et al.* *Nature* **502**, 532–535 (2013).
4. Birol, T., Benedek, N. A. & Fennie, C. J. *Phys. Rev. Lett.* **107**, 257602 (2011).
5. Zhu, X. *et al.* *J. Electr. Mater.* **32**, 1125–1133 (2003).

PALAEOLOGY

Inside–out turned upside–down

Sophisticated microscopy analysis of conodont elements suggests that these mysterious fossil structures are not, as has been previously suggested, evolutionary precursors to vertebrate teeth. [SEE LETTER P.546](#)

PHILIPPE JANVIER

Conodont elements are minute fossils resembling spines, combs or teeth that date from between 530 million and 200 million years ago. Like vertebrate teeth, scales and bones, they are formed of calcium phosphate, which makes them the earliest examples of mineralized, vertebrate-like skeletal structures. The fossils are believed to have belonged to early jawless vertebrates, and our understanding of the evolution of vertebrate anatomy is entwined in a long-standing, but debated, hypothesis that these structures were early homologues of teeth. But detailed comparisons of conodont elements of different ages, presented by Murdock *et al.*¹ in this issue (page 546), suggest that the sometimes striking structural resemblance between conodont elements and teeth is merely the result of evolutionary convergence*.

Conodont elements were discovered in 1856, and they have been widely used by geologists for dating marine sedimentary rocks because of their rapid shape changes over time. Indeed, the fossils were attributed to a wide range of animal (and even vegetal) groups until the discovery of the first complete 'conodont animal' in 1983. This fossil, called Conodontophorida (or 'conodont-bearing animal', thereafter called conodont), was preserved in 330-million-year-old rocks from

Scotland in the form of a soft-tissue imprint associated with an articulated assemblage of conodont elements². It showed an eel-shaped body with V-shaped muscle blocks, caudal-fin supports and presumed large eyes — features

suggesting that conodonts were closely related to, or members of, the vertebrates³.

The finding triggered extensive research comparing the tissue structure of conodont elements with that of teeth and odontodes — the minute denticles that can assemble to form scales, such as those covering the skin of sharks and rays. Odontodes derive from the vertebrate dermal (skin) tissue in modern vertebrates, but their similarities to the apparently internal conodont elements led to the 'inside-out' hypothesis, which suggests that odontodes arose from endodermal tissue in the vertebrate pharynx and then extended to the outer surface of the body in the form of scales. An extension of this hypothesis is that teeth arose before jaws. The topic has been the subject of extensive debate between researchers who support^{4–6} the notion that conodont

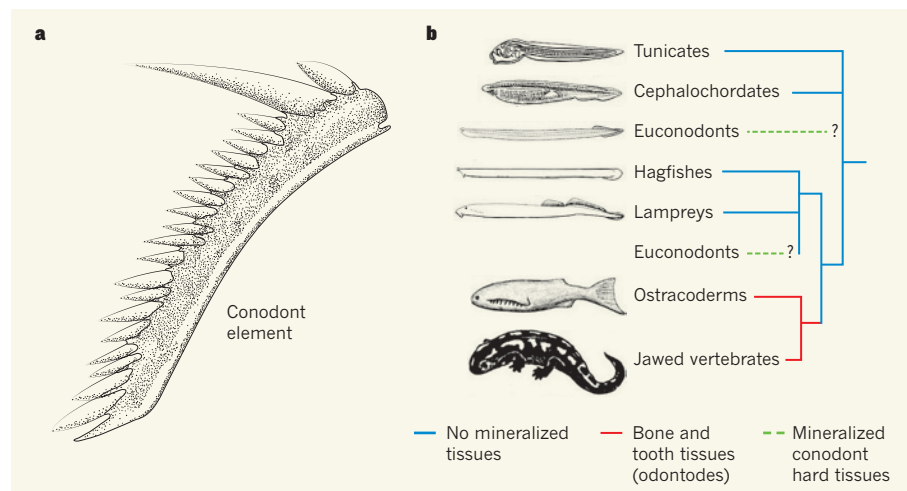


Figure 1 | False teeth. **a**, Some palaeontologists have considered the minute fossils called conodont elements to be homologues of the odontodes and teeth of vertebrates. **b**, Because of their ability to produce these mineralized skeletal structures, conodont animals were proposed^{3,5} to be a sister group of vertebrates in which bone and tooth tissues are present. But Murdock and colleagues' detailed analysis¹ suggests that the mineralized structures found in the later euconodont elements arose through a progressive assembly of hard tissues from earlier types of conodont element, which were very different from odontodes. This implies that euconodont elements arose independently (convergently) of vertebrate odontodes and teeth, and that the most recent common ancestor of conodonts and odontode-bearing vertebrates lacked mineralized skeletal tissues. This reasoning could position the divergence of euconodonts in the evolutionary branch of the cyclostomes (hagfishes and lampreys), or even earlier. (Drawings in **b** are taken from ref. 12.)

*This article and the paper under discussion¹ were published online on 16 October 2013.

elements and teeth are homologues and others who strongly reject⁷ this notion.

There are three main types of conodont element: protoconodonts, paraconodonts and euconodonts, approximately in order of age. The earliest protoconodont elements are simple, conical, hollow structures that are often poorly mineralized and that have been convincingly reinterpreted⁸ as grasping organs of fossil arrow worms (Chaetognatha), a group not closely related to vertebrates. Murdock and colleagues analysed conodont elements from the paraconodonts and euconodonts using high-resolution synchrotron radiation X-ray tomographic microscopy. The elements of paraconodonts and euconodonts are made of calcium phosphate, and Murdock *et al.* show that they display common features in their modes of growth. Paraconodont elements show a graded series of forms with ever-more-complex modes of growth, approaching that of euconodonts. Euconodont elements display a crown of enamel-like tissue that reinforces their resemblance to vertebrate odontodes. The finding that these mineralized structures arose in a step-wise manner from simple structures that are quite different from odontodes leads the authors to propose that euconodont elements evolved independently of vertebrate teeth and odontodes (Fig. 1).

Murdock and colleagues' dismissal of the previous suggestion of homology between these structures is a particularly brave conclusion because at least one of the authors has long been a strong supporter of the opposite view. Their findings demonstrate that the 'inside-out' theory is baseless and that the teeth and pharyngeal denticles of vertebrates are instead the consequence of the progressive invasion of the mouth and pharynx by skin tissues. But they also bring into question the evolutionary placement of conodonts. Without the mineralized structures, only the preserved soft tissues of conodonts reveal characteristics that they might share uniquely with vertebrates.

But how reliable are interpretations of these features? There is no clear evidence that the large, paired, anterior spots found on conodonts were eyes, although they look similar to the eye imprints preserved in fossils that are those of undisputed soft-bodied, jawless vertebrates. The muscle-block impressions in conodonts are V-shaped, whereas those of vertebrates are generally W-shaped, but this difference can be readily explained by post-mortem decay and compression⁹. The median fin supports seen in fishes were present in conodonts and were probably cartilaginous. Some sceptics of conodonts' place in the vertebrate club argue⁷ that all vertebrates have fin rays derived from dermal tissues, but this argument does not hold because extant hagfishes and lampreys, which date to 300 million and 360 million years ago, respectively, lack

mineralized tissues but are unambiguously recognized as vertebrates.

When all of these aspects are considered, the body of conodonts agrees with what could be expected from exceptionally preserved, soft-bodied, jawless vertebrates. I nevertheless concede that conodonts show no clear evidence of gill bars or gill pouches, traces of which are generally conspicuous in fossil lampreys and even in the 530-million-year-old presumed vertebrate *Haikouichthys*¹⁰. I am not desperately trying to shoehorn conodonts (or, more strictly, euconodonts) among the vertebrates, but in the face of the soft-tissue data alone, and for want of unambiguous evidence, I still believe that it is where they can best be classified. It is tempting to imagine conodonts swarming in the seas 530 million to 200 million years ago as scavengers comparable to the living hagfishes, with a mineralized feeding apparatus somewhat similar to that of the hagfishes' keratinous 'teeth'¹¹.

Although we can now attach to euconodonts the image of an animal that looks more like living jawless vertebrates than anything else, conodont elements have raised more

controversies than most other fossils and will probably continue to do so. But until a consensus is reached about their position in the tree of life, at least they will continue to help geologists date rocks. ■

Philippe Janvier is at the *Muséum National d'Histoire Naturelle*, 75231 Paris Cedex 05, France.

e-mail: janvier@mnhn.fr

1. Murdock, D. J. E. *et al. Nature* **502**, 546–536 (2013).
2. Briggs, D. E. G., Clarkson, E. N. K. & Aldridge, R. J. *Lethaia* **16**, 1–14 (1983).
3. Janvier, P. *Nature* **374**, 761–762 (1995).
4. Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308–1311 (1992).
5. Donoghue, P. C. J., Forey, P. L. & Aldridge, R. J. *Biol. Rev. Camb. Phil. Soc.* **75**, 191–251 (2000).
6. Sansom, I. J., Smith, M. P. & Smith, M. M. *Nature* **368**, 591 (1994).
7. Turner, S. *et al. Geodiversitas* **32**, 545–594 (2010).
8. Szaniawski, H. J. *Paleontol.* **56**, 806–810 (1982).
9. Sansom, R. S., Gabbott, S. E. & Purnell, M. A. *Palaeontology* **56**, 457–474 (2013).
10. Shu, D.-G. *et al. Nature* **421**, 526–529 (2003).
11. Goudemand, N., Orchard, M. J., Urdy, S., Bucher, H. & Tafforeau, P. *Proc. Natl Acad. Sci. USA* **108**, 8720–8724 (2011).
12. Janvier, P. *C.R. Palevol.* **8**, 209–219 (2009).

INORGANIC CHEMISTRY

A reducing role for boron

Carbon monoxide molecules are typically coupled together using metal catalysts. The discovery that boron, a non-metal, mediates such a reaction is startling, and raises the prospect of potentially useful carbon-carbon bond-forming processes.

POLLY L. ARNOLD

Ask any chemist about the reactivity of boron, and they will say that it forms strong bonds to carbon and oxygen. But the element is not known for its capacity as a reducing agent. Your chemist friends would therefore be surprised to hear of Braunschweig and colleagues' report¹ in *Nature Chemistry*, which describes the ability of a rather unusual boron compound not only to react and reduce molecules of carbon monoxide, but also to couple them together.

Carbon monoxide (CO) is a particularly reactive molecule, and is used in industry on a million-tonne scale as a source of single carbon atoms. Some of its most useful reactions require catalysts based on transition-metal complexes. The metal centre in these catalysts binds CO through its carbon atom and facilitates the molecule's insertion into adjacent groups also bound to the metal, forging a new carbon-carbon (C–C) bond. The reactions depend on the degree of electron transfer in the bond

between the catalytic transition metal and CO. Electron donation from the metal to the CO (known as back-donation) activates the CO for reaction and weakens its carbon-oxygen bond. Such reactions are widely used to make C–C bonds between CO and other substrates, but in general cannot be used to form these bonds between CO molecules themselves.

Non-metallic elements such as boron do not tend to facilitate C–C bond-forming reactions. But the boron-containing compound used by Braunschweig and colleagues in their reactions is highly unusual: it is a diboryne, a beautifully simple molecule in which two boron atoms are connected through a triple bond² (Fig. 1). Each boron atom is bound to a bulky molecule known as an *N*-heterocyclic carbene, which is characterized by carbon atoms that carry no formal charge, but that can donate a pair of electrons to other molecules. The binding of these carbenes helps to stabilize the diboryne.

Braunschweig and co-workers find that either their diboryne binds one CO molecule so that its carbon atom bridges the two boron

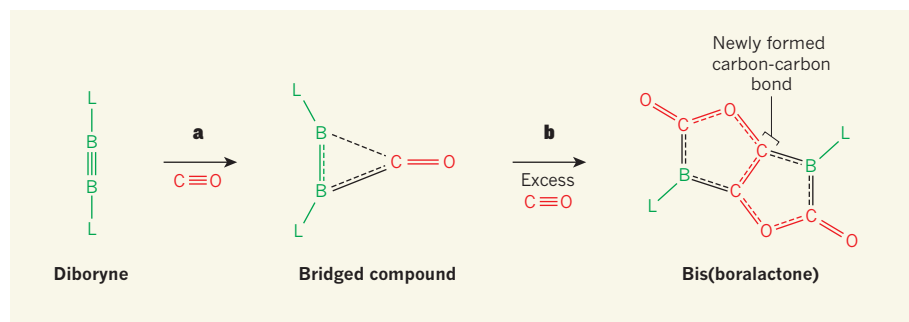


Figure 1 | Carbon monoxide molecules joined by a boron-containing compound. **a**, Braunschweig *et al.*¹ report that a single molecule of carbon monoxide (red) can be added to a diboryne compound (green) by bridging the diboryne's boron–boron triple bond. L is an *N*-heterocyclic carbene molecule. Dashed lines indicate partial bond formation. **b**, If an excess amount of CO (more than four equivalents) is added to the diboryne, four CO molecules are reduced and coupled together, forming a planar bis(boralactone) molecule that contains a new carbon–carbon bond. This is a rare example of a CO-coupling reaction mediated by a non-metal, and an equally rare example of a boron-mediated reduction reaction.

atoms (Fig. 1a), or that it couples four CO molecules together in a reduction reaction. The latter reaction yields a planar product called a bis(boralactone), in which the boron–boron triple bond is ruptured and a new C–C bond is formed (Fig. 1b). The diboryne binds and activates the first CO molecule through a back-donation of electron density similar to that seen in metal–CO compounds.

Diatomic boron units were originally observed³ in spectroscopy experiments as thermally unstable adducts containing two CO molecules, one bound at either end of the diboryne. In that case, the CO molecules donated two electrons each to the diboryne. By contrast, a single CO molecule binds as an electron acceptor to the thermally stable diboryne studied by Braunschweig *et al.*, bridging the two borons of the diboryne because the carbenes block access to the diboryne's ends. It will be fascinating to learn precisely how the diboryne back-donates electrons to activate the CO, because it does not possess the orbitals used by metal cations for back-donation.

Carbon monoxide can be induced to couple to other CO molecules if it is provided with electrons, allowing up to six molecules to react⁴. Coupling at the carbon atoms is the most common and desirable result⁵, but other bonding patterns are also possible. Braunschweig and colleagues observed that their diboryne can reduce and couple four CO molecules using a total of six electrons, three from each boron atom, in a rare example of a reduction by a boron compound. Intriguingly, all of the bonds in the B–C–C–B system of the product are shorter than equivalent single carbon–boron and C–C bonds, indicating that there is substantial electron delocalization in this molecule.

Braunschweig *et al.* describe the reducing power of their diboryne as strong: its reduction potential is –1.2 volts compared with that of ferrocene, a compound used as a standard for reduction potential. But this

is a small value relative to those for uranium systems that also reductively couple CO; all of these have potentials in solution⁶ of about –2 V, much more strongly reducing than ferrocene. It is exciting to imagine what further reduction chemistry might be possible with the diboryne, and whether simple changes to its carbenes could allow other compounds, formed from different numbers of CO molecules, to be isolated.

The real power of any new reaction for CO reduction depends on whether useful products containing more than one carbon atom can be made. For commercially viable applications, this will undoubtedly require the CO-coupled products to be reacted with hydrogen. Sadly, for the known metal-mediated CO-coupling systems, the only reaction with hydrogen so far reported occurs before any C–C bond formation, and so the resulting

hydrogenation–reaction products contain only one carbon atom⁷.

In another coup for boron, it was demonstrated⁸ earlier this year that electron-deficient boron-containing compounds known as boranes can activate both CO molecules and hydrogen (in combination with a suitable base), allowing CO to react with the hydrogen. It will therefore be interesting to see whether Braunschweig and colleagues' bis(boralactone), or analogues of it, will also react with hydrogen, perhaps when pre-activated by one of these electron-deficient boranes. If so, then the crucial question is whether the carbon–carbon bond within it remains intact — that is, whether a product containing two carbon atoms is formed. Also unknown is whether the authors' diboryne reagent can be easily recycled after reaction, a feature that could open up a new area of boron-catalysed chemistry. ■

Polly L. Arnold is at the EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ, UK.
e-mail: polly.arnold@ed.ac.uk

1. Braunschweig, H. *et al.* *Nature Chem.* <http://dx.doi.org/10.1038/nchem.1778> (2013).
2. Braunschweig, H. *et al.* *Science* **336**, 1420–1422 (2012).
3. Zhou, M. *et al.* *J. Am. Chem. Soc.* **124**, 12936–12937 (2002).
4. Watanabe, T., Ishida, Y., Matsuo, T. & Kawaguchi, H. *J. Am. Chem. Soc.* **131**, 3474–3475 (2009).
5. Summerscales, O. T., Cloke, F. G. N., Hitchcock, P. B., Green, J. C. & Hasari, N. *Science* **311**, 829–831 (2006).
6. Arnold, P. L. *Chem. Commun.* **47**, 9005–9010 (2011).
7. Frey, A. S. P., Cloke, F. G. N., Coles, M. P., Maron, L. & Davin, T. *Angew. Chem. Int. Edn* **50**, 6881–6883 (2011).
8. Dobrovetsky, R. & Stephan, D. W. *J. Am. Chem. Soc.* **135**, 4974–4977 (2013).

ASTRONOMY

New distance record for galaxies

Spectroscopic measurements of 43 candidates for distant galaxies have confirmed one to be the most remote galaxy securely identified to date — and it forms stars more than 100 times faster than the Milky Way. SEE LETTER P.524

DOMINIK A. RIECHERS

Light emitted by stars in far-off galaxies travels at a finite velocity and thus reaches us with a time delay, allowing us to probe deeper into the Universe's past with each more-distant object we find. The search for objects at greater distances from Earth than those already known is therefore important to improve our understanding of the Universe's history, and

necessary to eventually find the first generation of galaxies that formed after the Big Bang. These first galaxies are probably largely responsible for a major event in cosmic history: the 'reionization' of the neutral intergalactic hydrogen gas that filled space at these early epochs¹. In the early years of the past decade, astronomers successfully extended the distance over which we can observe galaxies time and time again, until progress stalled owing to technical limitations².

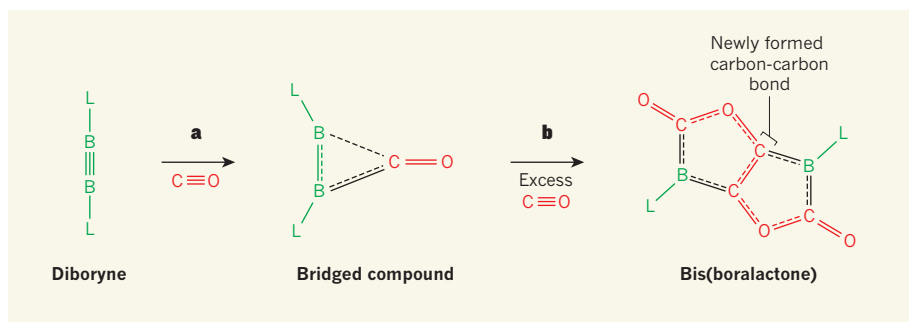


Figure 1 | Carbon monoxide molecules joined by a boron-containing compound. **a**, Braunschweig *et al.*¹ report that a single molecule of carbon monoxide (red) can be added to a diboryne compound (green) by bridging the diboryne's boron–boron triple bond. L is an *N*-heterocyclic carbene molecule. Dashed lines indicate partial bond formation. **b**, If an excess amount of CO (more than four equivalents) is added to the diboryne, four CO molecules are reduced and coupled together, forming a planar bis(boralactone) molecule that contains a new carbon–carbon bond. This is a rare example of a CO-coupling reaction mediated by a non-metal, and an equally rare example of a boron-mediated reduction reaction.

atoms (Fig. 1a), or that it couples four CO molecules together in a reduction reaction. The latter reaction yields a planar product called a bis(boralactone), in which the boron–boron triple bond is ruptured and a new C–C bond is formed (Fig. 1b). The diboryne binds and activates the first CO molecule through a back-donation of electron density similar to that seen in metal–CO compounds.

Diatomic boron units were originally observed³ in spectroscopy experiments as thermally unstable adducts containing two CO molecules, one bound at either end of the diboryne. In that case, the CO molecules donated two electrons each to the diboryne. By contrast, a single CO molecule binds as an electron acceptor to the thermally stable diboryne studied by Braunschweig *et al.*, bridging the two borons of the diboryne because the carbenes block access to the diboryne's ends. It will be fascinating to learn precisely how the diboryne back-donates electrons to activate the CO, because it does not possess the orbitals used by metal cations for back-donation.

Carbon monoxide can be induced to couple to other CO molecules if it is provided with electrons, allowing up to six molecules to react⁴. Coupling at the carbon atoms is the most common and desirable result⁵, but other bonding patterns are also possible. Braunschweig and colleagues observed that their diboryne can reduce and couple four CO molecules using a total of six electrons, three from each boron atom, in a rare example of a reduction by a boron compound. Intriguingly, all of the bonds in the B–C–B system of the product are shorter than equivalent single carbon–boron and C–C bonds, indicating that there is substantial electron delocalization in this molecule.

Braunschweig *et al.* describe the reducing power of their diboryne as strong: its reduction potential is –1.2 volts compared with that of ferrocene, a compound used as a standard for reduction potential. But this

is a small value relative to those for uranium systems that also reductively couple CO; all of these have potentials in solution⁶ of about –2 V, much more strongly reducing than ferrocene. It is exciting to imagine what further reduction chemistry might be possible with the diboryne, and whether simple changes to its carbenes could allow other compounds, formed from different numbers of CO molecules, to be isolated.

The real power of any new reaction for CO reduction depends on whether useful products containing more than one carbon atom can be made. For commercially viable applications, this will undoubtedly require the CO-coupled products to be reacted with hydrogen. Sadly, for the known metal-mediated CO-coupling systems, the only reaction with hydrogen so far reported occurs before any C–C bond formation, and so the resulting

hydrogenation–reaction products contain only one carbon atom⁷.

In another coup for boron, it was demonstrated⁸ earlier this year that electron-deficient boron-containing compounds known as boranes can activate both CO molecules and hydrogen (in combination with a suitable base), allowing CO to react with the hydrogen. It will therefore be interesting to see whether Braunschweig and colleagues' bis(boralactone), or analogues of it, will also react with hydrogen, perhaps when pre-activated by one of these electron-deficient boranes. If so, then the crucial question is whether the carbon–carbon bond within it remains intact — that is, whether a product containing two carbon atoms is formed. Also unknown is whether the authors' diboryne reagent can be easily recycled after reaction, a feature that could open up a new area of boron-catalysed chemistry. ■

Polly L. Arnold is at the EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ, UK.
e-mail: polly.arnold@ed.ac.uk

1. Braunschweig, H. *et al.* *Nature Chem.* <http://dx.doi.org/10.1038/nchem.1778> (2013).
2. Braunschweig, H. *et al.* *Science* **336**, 1420–1422 (2012).
3. Zhou, M. *et al.* *J. Am. Chem. Soc.* **124**, 12936–12937 (2002).
4. Watanabe, T., Ishida, Y., Matsuo, T. & Kawaguchi, H. *J. Am. Chem. Soc.* **131**, 3474–3475 (2009).
5. Summerscales, O. T., Cloke, F. G. N., Hitchcock, P. B., Green, J. C. & Hasari, N. *Science* **311**, 829–831 (2006).
6. Arnold, P. L. *Chem. Commun.* **47**, 9005–9010 (2011).
7. Frey, A. S. P., Cloke, F. G. N., Coles, M. P., Maron, L. & Davin, T. *Angew. Chem. Int. Edn* **50**, 6881–6883 (2011).
8. Dobrovetsky, R. & Stephan, D. W. *J. Am. Chem. Soc.* **135**, 4974–4977 (2013).

ASTRONOMY

New distance record for galaxies

Spectroscopic measurements of 43 candidates for distant galaxies have confirmed one to be the most remote galaxy securely identified to date — and it forms stars more than 100 times faster than the Milky Way. SEE LETTER P.524

DOMINIK A. RIECHERS

Light emitted by stars in far-off galaxies travels at a finite velocity and thus reaches us with a time delay, allowing us to probe deeper into the Universe's past with each more-distant object we find. The search for objects at greater distances from Earth than those already known is therefore important to improve our understanding of the Universe's history, and

necessary to eventually find the first generation of galaxies that formed after the Big Bang. These first galaxies are probably largely responsible for a major event in cosmic history: the 'reionization' of the neutral intergalactic hydrogen gas that filled space at these early epochs¹. In the early years of the past decade, astronomers successfully extended the distance over which we can observe galaxies time and time again, until progress stalled owing to technical limitations².

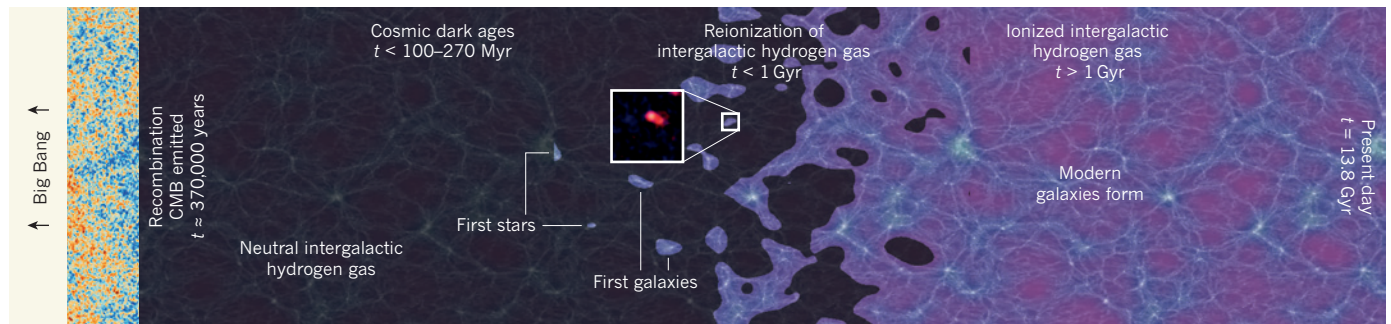


Figure 1 | Cosmic history and the first galaxies. Approximately 370,000 years after cosmic expansion started as a result of the Big Bang, the Universe had cooled sufficiently for protons and electrons to bind together (recombine) to form neutral hydrogen gas. At this time, the cosmic microwave background (CMB) radiation¹⁰ was emitted, and the Universe became opaque to hydrogen Lyman- α photons. These cosmic dark ages, an epoch that occurred between recombination and up to 270 million years (Myr) after the beginning of cosmic time, were ended by the first stars

and galaxies, which, once formed, reionized the Universe, allowing Lyman- α radiation to propagate freely. Thus, since this epoch of reionization, which had mostly been completed by 1 billion years (Gyr) after the Big Bang, Lyman- α emission from galaxies can be used to probe galaxy evolution and cosmic structure formation. Finkelstein *et al.*³ used Lyman- α emission to discover a galaxy at a cosmic age of just 700 million years (red galaxy shown in inset; image is about 10,000 parsecs across), probing deep into the epoch of reionization.

On page 524 of this issue, Finkelstein *et al.*³ present the discovery of the most distant galaxy found so far, observed at an epoch only 700 million years after the Big Bang (Fig. 1).

Owing to the expansion of the Universe that occurs as time passes, the wavelength of light emitted from galaxies far away undergoes redshift on its way to Earth — an effect that provides a direct measurement of distance. However, this process also results in key spectral features, such as the (typically brightest) Lyman- α line of hydrogen, eventually getting redshifted out of the visible-light range, which renders spectroscopic identification of the most distant galaxies challenging. Thus, despite the identification of dozens of strong candidates for very-high-redshift galaxies in deep imaging studies conducted⁴ with the infrared-sensitive Wide Field Camera 3 on the Hubble Space Telescope, progress in the spectroscopic confirmation of such candidates has slowed significantly in recent years.

A new generation of wide-field infrared cameras such as MOSFIRE at the W. M. Keck Observatory promises to remedy this situation. MOSFIRE can take sensitive spectra redwards of the visible-wavelength regime, typically for dozens of galaxies at a time. Finkelstein *et al.* used MOSFIRE to study 43 high-redshift galaxy candidates found in data from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) obtained with the Hubble Space Telescope⁵. They successfully identified redshifted Lyman- α line emission in a single galaxy, dubbed z8_GND_5296, setting a new distance record. The existence of only one more-distant object is securely known, owing to the explosion of a massive star about 70 million years earlier⁶. However, the galaxy associated with this event has remained undetected⁷. Remarkably, z8_GND_5296 actively forms stars at a rate more than 100 times higher than the Milky Way, considerably exceeding the star-formation activity of other galaxies at comparable distances. Indeed, galaxies

making stars as actively as z8_GND_5296 may represent the progenitors of the most extreme star-forming environments at later epochs than that of z8_GND_5296 (ref. 8).

Finkelstein *et al.* did not detect any Lyman- α line emission at comparable distances to z8_GND_5296 among the other 42 galaxy candidates — a success rate about six times lower than they anticipated. The authors consider it unlikely that the lack of additional identifications is caused by either the current technical limitations or the majority of the other galaxies being at less-extreme distances. On the basis of the unusual properties of the galaxy that they did identify, they discuss the hypothesis that the low confirmation rate could be due to either an unexpectedly low fraction of Lyman- α line emission escaping the galaxies or scattering of an unexpectedly high fraction of escaping Lyman- α line emission by significant amounts of neutral gas along most lines of sight.

Both possibilities could have far-reaching implications. The first possibility may indicate that galaxies at early epochs accrete gas at high rates, and that the resulting large amounts of gas extinguish most of the Lyman- α radiation in these young galaxies. The second possibility may suggest that cosmic reionization of the neutral intergalactic hydrogen gas has not progressed as far at the epoch of z8_GND_5296 as expected from other measurements⁹.

Clearly, Finkelstein and colleagues' study strongly motivates searches for other galaxies at the earliest epochs. It also highlights some of the challenges in identifying such galaxies from spectroscopy of a single emission line obtained with ground-based telescopes and from imaging at the galaxies' rest-frame ultraviolet and visible wavelengths, even when using the best facilities and highest-quality data available. The study further shows that even galaxies observed at a time when the Universe had reached only 5% of its current age may already be chemically enriched with dust and heavy elements (those heavier than

hydrogen and helium), which must have been produced by an earlier generation of stars. Heavy elements such as carbon, nitrogen and oxygen produce strong emission lines, which the James Webb Space Telescope (JWST) will be able to detect with relative ease in galaxies such as z8_GND_5296, after its launch towards the end of the decade. Such observations will remove the remaining ambiguities from the most challenging galaxy redshift measurements that are currently possible, and will provide significantly greater insight into the physical properties under which star formation takes place in these systems.

Even before the launch of the JWST, the Atacama Large Millimeter/submillimeter Array (ALMA) will provide the first substantial constraints on chemical enrichment of the first generation of galaxies based on far-infrared observations of lines from the same heavy elements and dust. The CCAT submillimeter survey telescope, anticipated to come online in 2018, will provide complementary samples of very-high-redshift galaxies selected directly by their dust content. Therefore, there is a bright future for studies of the first galaxies in the Universe. ■

Dominik A. Riechers is in the Department of Astronomy, Cornell University, Ithaca, New York 14853, USA.
e-mail: riechers@cornell.edu

- Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J. & Stark, D. P. *Nature* **468**, 49–55 (2010).
- Iye, M. *et al.* *Nature* **443**, 186–188 (2006).
- Finkelstein, S. L. *et al.* *Nature* **502**, 524–527 (2013).
- Bouwens, R. J. *et al.* *Nature* **469**, 504–507 (2011).
- Grogin, N. A. *et al.* *Astrophys. J. Suppl. Ser.* **197**, 35 (2011).
- Tanvir, N. R. *et al.* *Nature* **461**, 1254–1257 (2009).
- Chary, R., Surace, J., Carey, S., Berger, E. & Fazio, G. *GRB Coordinates Network, Circular Service* 9582, 1 (2009).
- Riechers, D. A. *et al.* *Nature* **496**, 329–333 (2013).
- Bolton, J. S. *et al.* *Mon. Not. R. Astron. Soc.* **416**, L70–L74 (2011).
- Planck Collaboration. Preprint at <http://arxiv.org/abs/1303.5062> (2013).



Cover illustration

Illustration by
Nik Spencer (concept:
Effie Apostolou and
John Bardakos)

Editor, *Nature*

Philip Campbell

Publishing

Richard Hughes

Production Editor

Jenny Rooke

Art Editor

Nik Spencer

Sponsorship

Reya Silao

Production

Ian Pope

Marketing

Elena Woodstock

Steven Hurst

Editorial Assistant

Abbie Williams

Anastasia Panoutsou

Transcriptional programs drive cell identity, dictating the specialized phenotypes of differentiated cells during the development of multicellular organisms. Interplay between signal-transduction pathways, transcription factors and the chromatin packaging of the genome sets the gene expression pattern of a cell, which must be relatively stably maintained once an organism is fully developed. The Reviews in this Insight explore how transcriptional states are regulated during development and disease.

The discovery that differentiated cells can be artificially reprogrammed into induced pluripotent stem cells by a small set of transcription factors has opened up exciting medical prospects and provided an opportunity to investigate how stable epigenetic states are built and reversed. Effie Apostolou and Konrad Hochedlinger discuss transcriptional and chromatin-based mechanisms behind cellular reprogramming and draw comparisons with tumorigenesis.

DNA methylation is a relatively stable epigenetic mark that locks genes in a silenced state, but pathways involved in removing DNA methylation marks are now becoming clear. Mechanistic models for these demethylation pathways are presented by Rahul Kohli and Yi Zhang, who highlight a key role for TET enzymes in various biological processes.

Transcriptional states become perturbed in disease, and genes encoding chromatin-associated proteins are often aberrantly expressed or mutated in cancer. Kristian Helin and Dashyant Dhanak discuss the latest discoveries of specific chemical inhibitors directed against chromatin regulators, as clinical trials of these compounds begin.

The transcriptional program of differentiated cells retains plasticity to ensure appropriate responses to the cellular environment. Philipp Gut and Eric Verdin propose that the metabolic state of a cell can directly influence DNA and histone modifications and thus help to regulate gene expression.

Finally, Wouter de Laat and Denis Duboule discuss the DNA regulatory elements termed enhancers that are at the heart of developmental transcription regulation and ensure the correct spatio-temporal expression of genes. Recent insight into the three-dimensional topology of chromosomes adds to our understanding of how enhancers act through long-range regulatory contacts.

Alex Eccleston, Francesca Cesari & Magdalena Skipper

Senior Editors

CONTENTS

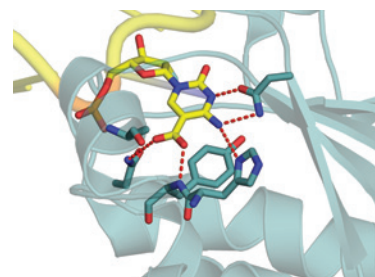
REVIEWS

462 Chromatin dynamics during cellular reprogramming

Effie Apostolou & Konrad Hochedlinger

472 TET enzymes, TDG and the dynamics of DNA demethylation

Rahul M. Kohli & Yi Zhang

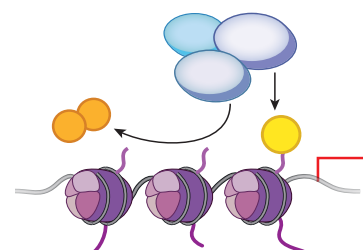


480 Chromatin proteins and modifications as drug targets

Kristian Helin & Dashyant Dhanak

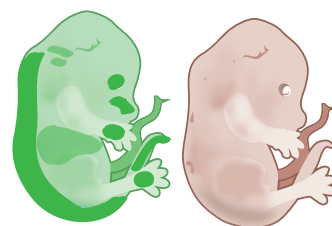
489 The nexus of chromatin regulation and intermediary metabolism

Philipp Gut & Eric Verdin



499 Topology of mammalian developmental enhancers and their regulatory landscapes

Wouter de Laat & Denis Duboule



Chromatin dynamics during cellular reprogramming

Effie Apostolou^{1,2,3,4} & Konrad Hochedlinger^{1,2,3,4}

Induced pluripotency is a powerful tool to derive patient-specific stem cells. In addition, it provides a unique assay to study the interplay between transcription factors and chromatin structure. Here, we review the latest insights into chromatin dynamics that are inherent to induced pluripotency. Moreover, we compare and contrast these events with other physiological and pathological processes that involve changes in chromatin and cell state, including germ cell maturation and tumorigenesis. We propose that an integrated view of these seemingly diverse processes could provide mechanistic insights into cell fate transitions in general and might lead to new approaches in regenerative medicine and cancer treatment.

Reprogramming of somatic cells to pluripotency can be achieved by different approaches, including somatic cell nuclear transfer (SCNT) into oocytes, fusion between somatic and pluripotent cells and ectopic expression of defined transcription factors^{1,2}. SCNT has demonstrated that epigenetic rather than genetic changes are the basis for most differentiation processes during normal development. Cell fusion experiments have documented that the pluripotent state is dominant over the somatic state in the context of hybrids. Together, these observations led to the seminal discovery that a small set of transcription factors, such as Oct4, Klf4, Sox2 and c-Myc (collectively called OKSM), are sufficient to convert differentiated cells into induced pluripotent stem cells (iPSCs)³. Importantly, induced pluripotency provides a biochemically and genetically tractable system to dissect the mechanisms underlying this remarkable cell fate change.

Recent progress in genome-wide technologies and the analysis of small cell numbers has allowed researchers to capture transcriptional and epigenetic snapshots of rare cell populations undergoing cell fate transitions in different biological contexts. These analyses have yielded important insights into the type and sequence of molecular changes inherent to transcription-factor-induced pluripotency, germ-cell reprogramming and cellular transformation. A common theme emerging from these studies is that nascent iPSCs, developing germ cells and pre-malignant cells use different as well as overlapping mechanisms to alter cell identity. The aim of this Review is to define those transcriptional, chromatin and epigenetic changes that endow specialized cells with pluripotency as well as the molecular barriers that resist cell fate change.

Mechanisms of induced pluripotency

Acquisition of induced pluripotency is a slow (around 2 weeks) and inefficient (0.1–3%) process^{1,3}, indicating that transcription factors need to overcome a series of epigenetic barriers that have been gradually imposed on the genome during differentiation to stabilize cell identity and to prevent aberrant cell fate changes. Earlier work has shown that cell populations expressing OKSM pass through a sequence of distinct molecular and cellular events (Fig. 1). Fibroblasts initially downregulate markers associated with the somatic state and subsequently activate genes associated with pluripotency, suggesting an ordered process^{4,5}. As soon as nascent iPSCs activate endogenous core pluripotency genes including *Oct4*, *Sox2* and *Nanog*, they acquire a self-sustaining pluripotent state and no longer require exogenous factor expression. The latter

events also coincide with the activation of the silenced X chromosome in female somatic cells, the upregulation of telomerase and the acquisition of immortality, which are hallmarks of cultured pluripotent cells^{5,6}. In the following sections, we will review our current knowledge of the way in which overexpressed transcription factors engage with chromatin, collaborate with epigenetic regulators and integrate extracellular signals to reprogram cellular identity (Fig. 2).

Transcription factors drive cell fate change

The most commonly used combination of reprogramming transcription factors comprises OKSM, and we will therefore primarily focus on this set of factors³. Earlier results have suggested that c-Myc and OKS play distinct parts during the acquisition and maintenance of pluripotency⁷. Briefly, OKS is the minimal set of factors required for iPSC generation from many cell types under classic reprogramming conditions (in the presence of serum and the cytokine leukaemia inhibitory factor, LIF). OKS cooperatively suppress lineage-specific genes and activate embryonic stem (ES)-cell-related genes, resulting in the establishment of a self-sustaining pluripotency network⁷. By contrast, ectopic c-Myc expression significantly enhances and accelerates reprogramming but is dispensable for iPSC formation^{8,9}. c-Myc expression functions early during reprogramming, presumably by stimulating cell proliferation and inducing a metabolic switch from an oxidative to a glycolytic state that is typical of pluripotent cells^{10,11}. More recent evidence suggests that c-Myc may also contribute to reprogramming by inducing pause release and promoter reloading of RNA polymerase, leading to transcriptional amplification of target genes^{12,13}.

It is worth noting that each of the original four reprogramming factors has been functionally replaced by either related transcription factors, upstream epigenetic modifiers, microRNAs (miRNAs) or small compounds¹. Moreover, iPSCs have been derived with molecules that do not contain any of the original transcription factors^{14,15}, indicating a remarkable flexibility and redundancy among reprogramming factors (Fig. 2). For example, a recent report demonstrated that the core pluripotency factors Oct4 and Sox2 can be substituted for by early lineage specifiers such as Gata3 and Geminin¹⁶. These molecules have been associated with mesendodermal and ectodermal differentiation, respectively; they also mutually repress the other respective lineage program, suggesting that the suppression of these major differentiation pathways is sufficient to trigger iPSC induction. This idea is consistent with the observation

¹Massachusetts General Hospital Center for Regenerative Medicine, 185 Cambridge Street, Boston, Massachusetts 02114, USA. ²Harvard Stem Cell Institute, 1350 Massachusetts Avenue, Cambridge, Massachusetts 02138, USA. ³Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA. ⁴Department of Stem Cell and Regenerative Biology, Harvard University and Harvard Medical School, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

that many classic pluripotency factors, including Oct4, Sox2 and Nanog, participate in early lineage decisions and hence may also be considered to be lineage specifiers¹⁷. A prediction that follows from these results is that transcription factors that stabilize a defined somatic state should be inhibitory to OKSM-mediated reprogramming. Indeed, depletion of the B-cell-specifying transcription factor Pax5 or ectopic expression of its antagonist C/EBP- α allows the reprogramming of terminally differentiated B cells¹⁸. Conversely, forced expression of differentiation-associated transcription factors, in combination with OKSM, significantly impairs iPSC formation by sustaining a somatic gene expression program and preventing activation of pluripotency genes¹⁹. Together, these findings demonstrate that reprogramming transcription factors have to achieve two key tasks, namely the extinction of the somatic program and the induction of a stable pluripotent state typical of ES cells.

Different types of chromatin targets

Developmental progression from pluripotent stem cells through progenitors to terminally differentiated cells is accompanied by a gradual deposition of repressive histone marks, followed by chromatin compaction^{20–22}. A key question is therefore how reprogramming transcription factors dismantle somatic chromatin and establish an epigenetic state that is compatible with pluripotency. Recent studies assessing OKSM occupancy and histone marks early during the reprogramming of mouse and human fibroblasts into iPSCs have provided the first clues for answering this question^{10,23,24}. Combining these observations, one may categorize OKSM targets into three classes of loci based on chromatin accessibility, the requirement for additional remodelling and the kinetics of transcriptional activation (Fig. 3a). Genes with an 'open' chromatin state in somatic cells comprise the first group of targets, characterized by increased DNaseI hypersensitivity, active di- and tri-methylation of histone H3 lysine 4 (H3K4me2 and H3K4me3) and the ability to bind OKSM immediately. Downregulated somatic genes and genes linked to a mesenchymal-to-epithelial transition (MET), which specifies early stages of reprogramming²⁵, fall into this group²⁴.

A second class of early bound OKSM targets includes distal regulatory elements, which seem to require additional chromatin remodelling for transcriptional activation²⁴. A subgroup of these elements carries the H3K4me1 mark and exhibits nucleosome depletion as well as DNase I hypersensitivity, which are chromatin features characteristic of 'permissive enhancers'. Permissive enhancers typically bind transcription factors before their associated promoter regions and prior to transcriptional activation²⁶. The *MyoD* locus exemplifies this group of enhancers; ectopically expressed Oct4 initially binds to the *MyoD* enhancer, triggering crosstalk with its promoter and subsequent acquisition of a poised chromatin state²⁶. Another subset of distal regulatory elements comprises DNase-I-resistant loci that are unable to bind c-Myc alone²⁴. Early pluripotency genes such as *Sall4* belong to this group. Interestingly, occupancy of these targets by OKS facilitates binding of c-Myc. This observation thus identifies OKS as 'pioneer factors', defining the ability of transcription factors to bind closed somatic chromatin and allow chromatin remodelling as well as recruitment of other transcription factors and cofactors²⁴.

Broad heterochromatic regions enriched for the repressive H3K9me3 mark constitute a third set of OKSM targets. Genes within this category comprise core pluripotency genes, such as *Nanog* and *Sox2* (ref. 24). These regions are refractory to immediate OKSM binding and seem to require the most extensive chromatin remodelling for transcriptional activation. Of interest, refractory domains are reminiscent of broad chromatin regions enriched for the H3K9me2 mark, termed large organized chromatin K9-modification (LOCK)²⁷. LOCKs are generally underrepresented in pluripotent cells and associated with repression of lineage-specific genes in differentiated cells²⁷.

Bivalent genes constitute a separate group of chromatin targets that are marked by both active H3K4 methylation and repressive H3K27 methylation in iPSCs or ES cells²⁸. Genes within this category are transcriptionally silent in ES cells and iPSCs but poised for rapid activation

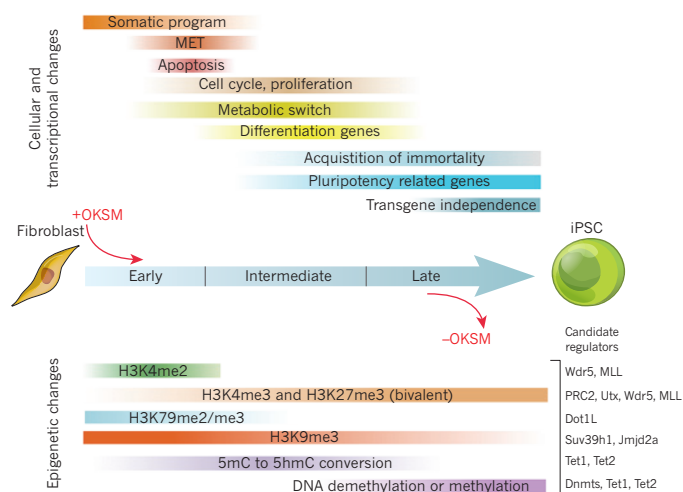


Figure 1 | Dynamics of key molecular events during direct reprogramming.

A summary of cellular, transcriptional and epigenetic changes (coloured bars) that occur during induced pluripotent stem cell (iPSC) formation from fibroblasts and examples of candidate regulators that have been associated with the depicted chromatin marks in the context of direct reprogramming (right). Red arrows indicate the time points of exogenous factor induction (+OKSM) and withdrawal (–OKSM). 5hmC, 5-hydroxymethylcytosine; 5mC, 5-methylcytosine; MET, mesenchymal-to-epithelial-transition; OKSM, Oct4, Klf4, Sox2, c-Myc.

on lineage commitment. Although the actual number and relevance of bivalent promoters remain controversial²⁹, induced pluripotency gradually restores these domains to an ES-cell-like pattern^{10,23}. The transcription factor Utl1 has recently been implicated in the regulation of bivalency by a dual mechanism that involves deposition of the repressive H3K27me3 mark and degradation of residual transcripts³⁰. Notably, Utl1 expression can substitute for some of the original reprogramming factors, providing indirect evidence that the establishment of bivalent promoters may be an important step during the acquisition of pluripotency¹⁴. Together, these results illustrate how the initial chromatin state of pluripotency-related and somatic genes determines if and when they become occupied and transcriptionally regulated by OKSM in the course of reprogramming. They further suggest that certain histone marks (for example, H3K9 methylation) act as potent barriers that resist acquisition of pluripotency. We will therefore next focus on the enzymes that deposit or remove these marks and their impact on cellular reprogramming.

Role of histone-modifying enzymes

Histone marks and chromatin structure are regulated by histone-modifying enzymes, or 'writers', such as histone methyltransferases (HMTs) and histone acetyltransferases (HATs), and 'erasers' such as histone demethylases (HDMs) and histone deacetylases (HDACs) (Fig. 2). These enzymes function as co-activators or co-repressors of OKSM at different stages of reprogramming and can profoundly influence iPSC derivation. For example, recruitment of polycomb repressive complex 2 (PRC2), which deposits the repressive H3K27me3 mark, and inhibition of Dot1L (which establishes the active H3K79me2 and H3K79me3 marks) have been associated with the downregulation of somatic genes early in reprogramming^{31,32}. Accordingly, loss of PRC2 abrogates whereas loss of Dot1L enhances iPSC formation. Activation of the H3K36 HDMs Jhdm1a and Jhdm1b³³ and suppression of the H3K27 HDM Jmjd3 promote intermediate-to-late stages of iPSC generation by suppressing the *Ink4/Arf* locus³⁴, which is essential for the acquisition of immortality. An additional early role for Jhdm1b in epithelial gene activation has recently been reported³⁵. By contrast, H3K9 HMTs maintain the above-mentioned 'refractory' heterochromatic state of somatic cells and thus act as major barriers of reprogramming. Consistent with this notion, knockdown of *G9a* (H3K9me2 HMT), or *Suv39h1* and *Suv39h2* and *Setdb1* (H3K9me3 HMTs), or overexpression of H3K9 HDMs,

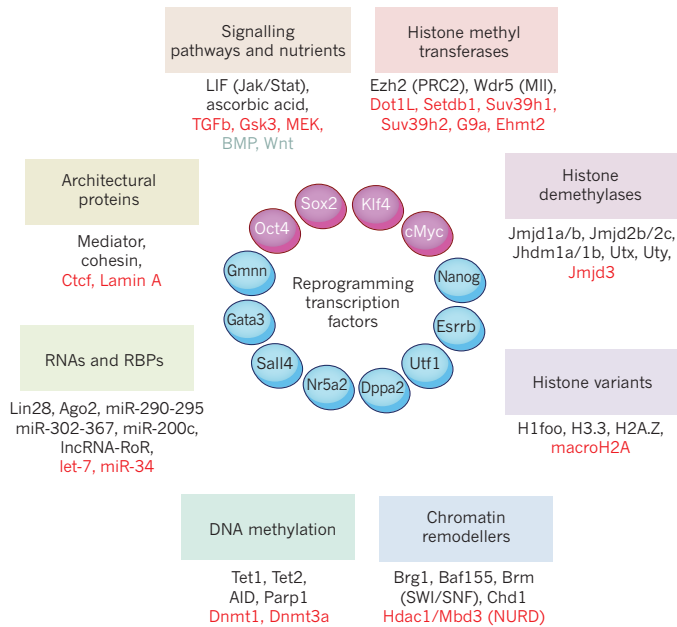


Figure 2 | Interplay between reprogramming factors and molecules influencing chromatin state. Different transcription factors that have been shown to trigger induced pluripotency, with the most broadly used combination (Oct4, Klf4, Sox2, c-Myc) highlighted. Some molecules have been shown to facilitate (black) or inhibit (red) reprogramming. Bone morphogenetic proteins (BMPs) and Wnts have stage-dependent enhancing or suppressive roles during iPSC formation. RBPs, RNA binding proteins.

increase transcription factor accessibility and result in more efficient iPSC generation from somatic cells^{24,36,37}. Altogether, these results demonstrate that histone code writers and erasers are essential components of iPSC formation by either maintaining the somatic state or assisting in the transcription-factor-induced establishment of pluripotency.

Reprogramming transcription factors have been reported to directly interact with histone-modifying enzymes, providing a mechanistic explanation for how they may alter chromatin and cell state during induced pluripotency. Examples include the H3K27 HDM Utx¹⁴ and the H3K4me3 'effector' Wdr5 (ref. 38) that bind to Oct4 protein and co-occupy many genomic targets in ES cells, thus keeping them in a transcriptionally active state. Depletion of either molecule blunts iPSC formation owing to a failure to activate pluripotency genes^{14,38}. Oct4 may have a very specific role in recruiting epigenetic regulators to target genes compared with Sox2, Klf4 and c-Myc, because it cannot be replaced by related family members during induced pluripotency⁸. In agreement, deletion of a linker domain on Oct4, which is absent on other POU domain family members and associates with chromatin remodellers implicated in reprogramming, abrogates iPSC formation³⁹. Intriguingly, some reprogramming-associated cofactors function in a chromatin-independent manner. For example, the H3K27 HDM Jmjd3 blocks reprogramming not only by activating the *Ink4a/Arf* locus but also by targeting the methyl-lysine effector protein Phf20 for ubiquitination; Phf20 is required to activate *Oct4* transcription in collaboration with Wdr5 (ref. 34). Overall, these and other examples document the physical association of reprogramming transcription factors with a variety of histone modifiers and exemplify the diverse mechanisms by which they assist in reinstating pluripotency in somatic cells.

DNA methylation safeguards cell identity

DNA methylation is considered to be the most stable epigenetic modification, which confers permanent gene silencing during development and in the adult. Changes in histone modifications typically precede the removal or deposition of DNA methylation marks during differentiation²⁰. Similarly, DNA methylation changes almost exclusively occur

at the end of the reprogramming process and after chromatin changes have taken place¹⁰, indicating a hierarchy of events that recapitulates normal development (Fig. 1). DNA methylation is established by the *de novo* methyltransferases Dnmt3a and Dnmt3b and preserved by the maintenance methyltransferase Dnmt1 (ref. 40). Although *DNMT3a* knockdown promotes iPSC formation in human cells³¹, deletion of the mouse enzymes Dnmt3a and Dnmt3b has no consequence for cellular reprogramming⁴¹. This surprising finding suggests that the silencing of lineage-specific genes is mainly achieved through alternative mechanisms such as deposition of repressive H3K27 methylation, which is consistent with the essential role of PRC2 in iPSC formation³².

In contrast to the dispensability of *de novo* methylation for iPSC formation, DNA demethylation of pluripotency genes seems to be crucial for faithful reprogramming. Demethylation can occur by either active or passive mechanisms (see Review by Kohli and Zhang⁴⁰), both of which have been implicated in iPSC generation. Downregulation of Dnmt1 in reprogramming intermediates facilitates their transition towards authentic iPSCs, consistent with the supportive role of passive, replication-dependent demethylation in iPSC formation⁴². Enzymes associated with active DNA demethylation have a more direct link with iPSC formation (Fig. 1). TET proteins catalyse the hydroxylation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), which serves as a substrate for TDG-mediated base excision repair into unmodified cytosine⁴⁰. Shortly after overexpression of OKSM, Tet2 induces hydroxymethylation of key pluripotency genes such as *Nanog* and *Esrrb*, priming them for subsequent demethylation and transcriptional activation (Figs 1 and 3b)⁴³. Interestingly, proteomic and genomic analyses revealed that Tet1 and Tet2 directly interact with Nanog and co-occupy many pluripotency targets in ES cells, implicating Nanog in the targeting of Tets⁴⁴. In agreement, simultaneous overexpression of Tet1 or Tet2 together with Nanog significantly enhances, whereas depletion of Nanog or Tet2 abrogates, iPSC formation^{43–45}. Moreover, Tet1 overexpression can compensate for exogenous Oct4 expression during cellular reprogramming, providing genetic evidence that Tet1 contributes to the activation of somatically silenced pluripotency genes⁴⁶. A role for activation-induced deaminase (Aid) in DNA demethylation during transcription-factor-induced reprogramming has also been reported, although the underlying mechanisms are incompletely understood⁴⁰.

Inefficient DNA demethylation or remethylation has further been suggested to be the main reason for the 'epigenetic memory' observed in many iPSC lines. This term describes cell-type-of-origin-dependent transcriptional and epigenetic patterns that can influence the differentiation potential of iPSCs¹. Of note, genomic regions that fail to undergo DNA demethylation towards an ES-cell-like state in human iPSCs were shown to be decorated by broad H3K9me3 domains in somatic cells⁴⁷, and some of these areas overlap with the above mentioned 'refractory domains' that are inaccessible to OKSM early in reprogramming²⁴. These findings therefore suggest that promiscuous OKSM binding or lack of OKSM binding to targets could explain some of the observed differences between ES cells and iPSCs. It will be interesting to assess whether manipulation of H3K9 HDMs or HMTs is sufficient to erase epigenetic memory in iPSCs.

Three-dimensional chromatin architecture in reprogramming

Accumulating evidence suggests that local and three-dimensional (3D) chromatin architecture provide additional levels of gene regulation in pluripotent stem cells (see Review by de Laat and Duboule⁴⁸). However, their roles in cellular reprogramming are incompletely understood. Local chromatin architecture defines the position and density of nucleosomes as well as the presence of histone variants (Figs 2 and 3c)⁴⁹. Histone variants usually modify the ability of nucleosomes to undergo remodelling and to accommodate active or repressive histone modifications. The histone variant macroH2A has previously been associated with resistance to efficient chromatin remodelling⁴⁹. In agreement, the presence of macroH2A potentially inhibits transcription-factor-induced reprogramming of somatic cells to pluripotency

by maintaining pluripotency loci in a repressed state^{50–52}.

Local chromatin architecture is regulated by diverse remodelling complexes, which also affect iPSC formation (Fig. 2). Components of the SWI–SNF complex, such as Brg1, Baf155 and Brm, are directly recruited by Oct4 to targets in order to relax chromatin structure and facilitate binding of other transcription factors⁵³. This finding thus corroborates that Oct4 has a role as a pioneer factor by influencing local chromatin structure at silenced targets. Similarly, the CHD remodelling factor Chd1 has been proposed to actively open chromatin during factor-induced reprogramming, and knockdown of its gene impairs iPSC formation⁵⁴. By contrast, members of the repressive NURD complex including Hdac1 and Mdb3, which are crucial for heterochromatin compaction, inhibit reprogramming and their gene knockdown strongly increases the efficiency of iPSC generation^{55,56}. Interestingly, Mdb3 associates with loci enriched for Tet1 and 5hmC in ES cells and its expression is essential for global levels of 5hmC⁵⁷. The latter observation may explain the discrepancy between the observed early deposition of 5hmC at pluripotency loci⁴³ but its delayed conversion into unmodified C during iPSC generation¹⁰; Mdb3 may be recruited to these hydroxymethylated pluripotency promoters and block immediate demethylation until unidentified co-activators relieve Mdb3-mediated gene repression.

In addition to local chromatin structure, 3D chromatin architecture has been implicated in pluripotency, differentiation and reprogramming (Fig. 3d)⁴⁸. Differentiation of ES cells is accompanied by repositioning of pluripotency genes from the nuclear centre to the nuclear periphery³⁸ and a disruption of promoter–enhancer looping at key pluripotency loci such as *Oct4* (refs 59, 60) and *Nanog*^{54,55}. This raises the question of whether and how 3D chromatin structure is restored in iPSCs. A recent study identified complex pluripotency specific long-range interactions of the *Nanog* locus, which become rearranged during differentiation and are largely restored during reprogramming⁶¹. The establishment and maintenance of this network is dependent on mediator and cohesin complexes, which are known to orchestrate long-range chromatin interactions⁶⁰. Interestingly, subunits of these complexes were found to directly interact with reprogramming factors^{61,62} and their knockdown inhibited iPSC formation⁶¹.

Extending these findings, long-range chromatin interactions around the *Oct4* promoter were recently implicated in the reprogramming of murine and human cells^{62,63}. Importantly, these interactions took place specifically in those rare cells that were poised to form iPSCs, and they preceded transcriptional activation, suggesting a causal effect for 3D chromatin structure on transcription. These results — together with previous studies documenting a role for Oct4 and Klf1 (highly similar to Klf4) in mediating long-range chromatin interactions^{64,65} — support the idea that reprogramming factors do not merely activate or silence genes but also function as chromatin organizers, which rearrange chromatin architecture from a somatic to a pluripotent state. Furthermore, this interpretation is consistent with the recent discovery of ‘super-enhancers’, which are broad distal regulatory elements characterized by cooperative and excessive binding of mediator components and cell-type-specific transcription factors, such as Esrrb and Klf4, in ES cells⁶⁶. Given the documented role of super-enhancers in controlling the expression of master regulatory genes in different cell types, it is likely that the resetting of somatic-specific to pluripotency-specific super-enhancers constitutes another roadblock for iPSC formation. The dynamics of this switch during induced pluripotency and the potential role of super-enhancers in 3D chromatin architecture certainly warrants further investigation.

Cell signalling and chromatin

External cues are crucial to direct cells expressing OKSM towards a stable pluripotent state and to prevent acquisition of alternative cell fates⁷. Extracellular signals can either support or inhibit iPSC formation (Fig. 2). For example, dual chemical inhibition of the Gsk3b, and Erk1 and Erk2 pathways (known as the 2i condition) enhances the transition

of partially reprogrammed cells into iPSCs⁶⁷. Similarly, activation of the Jak–Stat3 pathway by the cytokine LIF is limiting for iPSC formation⁶⁸. By contrast, the TGF- β pathway negatively affects iPSC generation and suppression of the pathway by chemical inhibitors significantly increases the acquisition of pluripotency^{69,70}.

Recent data offer new mechanistic insights into how external signals communicate with chromatin structure in pluripotency and reprogramming. For instance, the downstream effector of LIF signalling, Stat3, requires the chromatin remodeller subunit Brg1 to keep targets accessible and prevent their repression by the PRC2 complex⁷¹. Moreover, culture of ES cells in 2i endows cells with a so-called naive or ground state that is characterized by altered H3K27me3 distribution and a decreased number of bivalent promoters as well as global DNA hypomethylation^{72,73}. Mechanistically, growth of ES cells in 2i induces activation of the transcription factor Prdm14, which directly represses Dnmt3a and Dnmt3b and inhibits differentiation-inducing Fgfr signalling⁷². It is important to mention here that signalling molecules can also have opposite effects on different stages of reprogramming. Bone morphogenetic proteins promote an early MET in an miRNA-dependent manner²⁵, whereas they block the late conversion of partially reprogrammed cells into iPSCs by targeting the repressive H3K9 HMTs Suv39h1 and Setdb1 (ref. 36). Conversely, Wnt–Tcf3 signalling is inhibitory early but stimulatory late in reprogramming⁷⁴.

Nutrients and cofactors present in the extracellular environment represent a final class of molecules that influence the epigenome and cellular reprogramming. A case in point is ascorbic acid (vitamin C), which has

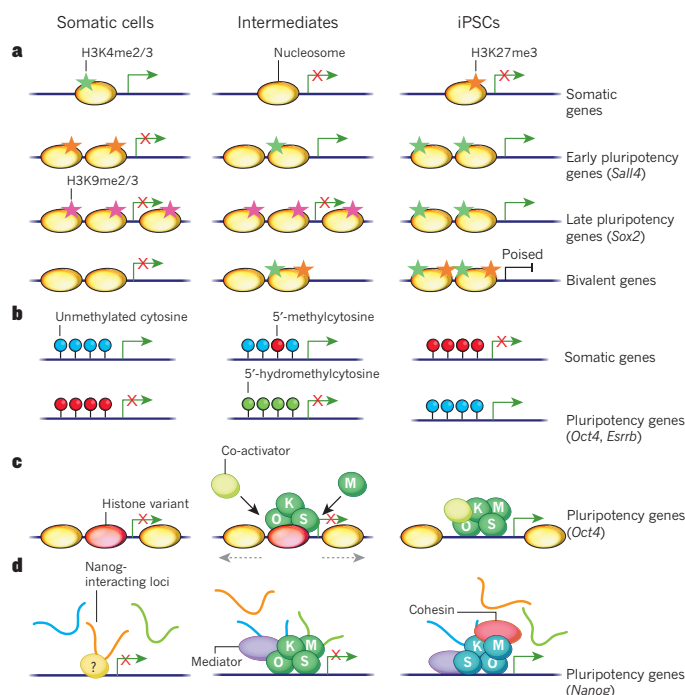


Figure 3 | Levels of epigenetic gene regulation during induced pluripotency. **a**, Four categories of genes with associated histone modifications and their transcriptional response during reprogramming. Examples of genes are shown in parentheses. **b**, Gain and loss of DNA methylation occurs late in reprogramming, whereas the acquisition of hydroxymethylation at pluripotency genes takes place at early-to-mid stages of iPSC formation. **c**, Oct4 (O), Klf4 (K) and Sox2 (S) function as ‘pioneer factors’ that bind to high-density nucleosome regions, allowing chromatin remodelling (indicated by dotted arrows) and the recruitment of other factors including c-Myc (M). **d**, Changes in long-range chromatin interactions around the *Nanog* locus (somatic-specific, orange loop; intermediate-specific, green; and pluripotency specific, blue interacting loci) during iPSC formation. iPSC formation re-establishes a 3D chromatin network typical of ES cells, and this process depends on mediator and cohesin.

been shown to strongly enhance the efficiency and kinetics of reprogramming⁷⁵ and to increase the quality of mouse iPSCs by preventing aberrant hypermethylation⁷⁶. Ascorbic acid presumably functions both as an antioxidant and as a cofactor for specific epigenetic modifiers such as the H3K36 HDMs Jmjd1a and Jmjd1b³³. Furthermore, ascorbic acid has been suggested to be a cofactor for H3K9 HDMs and Tet enzymes, according to recent studies that reported a global decrease of the repressive H3K9me2 and H3K9me3 marks³⁶ and genome-wide DNA hypomethylation⁷⁷, respectively, in nascent iPSCs exposed to this compound. Together, these observations provide compelling new evidence for the tight communication between reprogramming-associated signalling molecules and transcription factors in order to rewire epigenetic regulatory circuits.

Sequence of molecular events

An unresolved question is whether the rare cells that give rise to iPSCs undergo a defined sequence of transcriptional and epigenetic events that are essential for successful reprogramming. Different approaches have been used to resolve this issue. A number of reports used surface marker combinations to prospectively identify those rare cell populations that are destined to form iPSCs^{10,78–80}. One of these studies identified two major waves of gene expression changes that coincided with the early extinction of somatic genes and with the late activation of core pluripotency genes, respectively¹⁰. The lack of major transcriptional changes during the intermediate phase suggested that cells undergo gradual epigenetic alterations to prime the genome for transcriptional activation of pluripotency genes. In agreement with this is the observation that histone marks associated with pluripotency enhancers are established at early and intermediate stages of reprogramming, whereas DNA methylation changes occur late, coinciding with transcriptional activation of associated genes¹⁰. Integrating observations from other studies, this intermediate period is also characterized by the establishment of pluripotency-specific long-range chromatin interactions^{61,62} and Tet-mediated conversion of 5mC into 5hmC at pluripotency promoters⁴³, further supporting the interpretation that cells undergo a series of chromatin changes in preparation for stable pluripotency.

An independent study analysed transcriptional changes of 48 genes in single cells undergoing reprogramming¹⁵ and concluded that the initial response to reprogramming factor expression is quite heterogeneous and consistent with a stochastic process. However, later events, leading to the activation of pluripotency genes, occur in a hierarchical manner. This analysis led to the inference that activation of the endogenous *Sox2* locus is a crucial upstream event in cells that undergo successful reprogramming. A parallel study examined clonal populations of cells expressing OKSM and defined ‘maturation’ and ‘stabilization’ phases of reprogramming that were distinguished by differential expression of pluripotency genes⁸¹. Unexpectedly, the authors discovered that the transition to the stabilization phase is dependent on a different set of factors (for example, GDNF signalling and meiosis genes) from those that control maintenance of pluripotency.

Several groups have reported transient upregulation of developmental regulators, such as epidermal, extra-embryonic and epiblast-associated genes, at intermediate stages of reprogramming^{10,16,42,74,78} (Fig. 1). Although the molecular mechanisms underlying this observation remain elusive, it is tempting to speculate that reprogramming intermediates transiently pass through a state with increased developmental plasticity that could represent stages of normal development⁷. Alternatively, these genes might be activated as a consequence of aberrant transcription-factor binding¹¹ or unspecific effects incurred by small compounds⁸². Regardless, recent studies showed that depletion of some of these transiently expressed genes impairs reprogramming into iPSCs, suggesting functional relevance^{79,82}.

In conclusion, although the overall gene expression trends are similar among different studies, variability related to the exact sequence of molecular events and the relative contribution of stochastic and deterministic events remain to be resolved. Another fundamental question that needs

to be addressed is whether cells expressing alternative reprogramming factors pass through the same sequence of events described here, and encounter the same barriers as cells expressing OKSM. This question is particularly relevant with respect to reprogramming approaches that involve small compounds⁸² or transcription factors that have not been directly associated with the core pluripotency network¹⁶.

Lessons from other reprogramming systems

In this section, we will compare induced pluripotency with other processes involving epigenetic reprogramming such as SCNT, cell fusion and germ-line specification with the goal of identifying mechanistic similarities and differences (Table 1).

Somatic cell nuclear transfer

Remarkably, the somatic gene expression program is downregulated within 22 hours of SCNT in mice (D. Egli, personal communication)⁸³. This observation is reminiscent of iPSC formation and suggests that extinction of somatically expressed genes is rapid and efficient in both approaches. Somatic *Oct4* activation after SCNT occurs with similar kinetics (24–36 hours)⁸³ and has been suggested to require Tet3 (ref. 84), whereas it takes a minimum of 8 days to detect *Oct4* expression during induced pluripotency⁸³ and this process seems to involve Tet2 (ref. 43). This marked difference in pluripotency gene activation poses the key question of whether SCNT and induced pluripotency depend on the same transcription factors for reprogramming; both *Oct4* and *Sox2* are in fact detectable in oocytes. Schöler and colleagues recently addressed this question by genetically depleting maternal *Oct4* from mouse oocytes before SCNT⁸⁵. Surprisingly, loss of *Oct4* did not abrogate the oocyte's ability to reprogram somatic nuclei, indicating compensation by other factors. The identification of the oocyte-enriched Glis1 protein⁸⁶, capable of enhancing iPSC formation in the context of OKS expression, supports this hypothesis.

A key molecular event that distinguishes induced pluripotency from SCNT is the rapid histone exchange between somatic nucleus and oocyte, as demonstrated in *Xenopus* SCNT experiments. Specifically, the somatic linker histone H1 is replaced within hours of SCNT for the oocyte-specific counterpart B4, and this process is essential for pluripotency gene activation in reconstructed oocytes⁸⁷. This particular exchange of histone types might contribute to reprogramming by depleting somatic chromatin from epigenetic repressors known to interact with histone H1, such as Dnmt1 and Dnmt3b, and H3K9 methyltransferases^{88,89}. Concomitant with the replacement of ‘repressive’ histone variants, incorporation of ‘active’ histone variants including H3.3 and H2A.X into the somatic chromatin facilitates efficient chromatin remodelling of embryonic genes^{90,91}. Despite these effective mechanisms, some epigenetic marks including those of the silenced X chromosome (Xi) in female somatic nuclei seem to be more recalcitrant to remodelling by the oocyte compared with pluripotency genes, indicating differential susceptibility of some genomic loci to the oocyte's reprogramming machinery⁹². The relative resistance of X reactivation to efficient remodelling during *Xenopus* SCNT has been functionally linked to the repressive histone variant macroH2A, because its knockdown resulted in a more efficient reactivation of the Xi⁹². Thus, the eviction of macroH2A represents a rate-limiting step for successful reprogramming in the context of both SCNT and iPSC induction. Given the prominent role of ‘active’ and ‘repressive’ histone variants during SCNT, it should be informative to systematically test their function and that of associated chaperones⁴⁹ during iPSC generation.

Cell–cell fusion

Downregulation of somatic genes in ES-cell-somatic hybrids also occurs within the first 1–2 days of fusion⁹³. When examining the same *Oct4*–GFP reporter that was used for SCNT and iPSC formation, *Oct4* reactivation in fusion hybrids took place with similar kinetics as SCNT (24–48 hours)⁹³ and this process was reported to involve Tet2 (ref. 94).

Table 1 | Comparison of types of cellular reprogramming

Type of reprogramming	Extinction of cell-of-origin expression program (example)	Activation of pluripotency genes (example)	X chromosome reactivation (example)	Histone variants (chaperones)	Histone modifiers	DNA methylation factors	Requirement for endogenous transcription factors	Relevant signalling pathways
Direct reprogramming (OKSM factors)	1–2 d (<i>Thy1</i> and <i>Snai1</i> during fibroblast reprogramming) ⁵	8–12 d (Oct4–GFP reporter) ^{4,5}	Around 12 d (X–GFP reporter) ⁵	macroH2A ^{50–52}	PRC1 and PRC2 (ref. 32), Utx ¹⁴ , Dot1L ³¹ , HDACs ¹ , (see Fig. 1)	Dnmt1 (ref. 42) Tet1 and Tet2 (refs 43, 44) Aid ⁴⁰	Nanog ⁴⁵ , Stat3 (ref. 67)	Wnt ^{1,73} , LIF ⁶⁷ , BMP ^{25,36} , ERK or GSK3b inhibition ('2i') ⁶⁶
SCNT	22–24 h (2-cell embryo; global silencing of somatic genes and activation of zygotic genes) ⁸²	<36 h (4-cell embryo; Oct4–GFP reporter) ⁸²	Around 3 d (morula embryo; X–GFP reporter)	H1 or B4 (ref. 86), H3.3 (ref. 89), H2A ^{X90} , H2A ^{Z90} , macroH2A ⁹¹	Hdacs ⁹¹	Dnmt1 Tet3 (ref. 83)	Oct4 not required ⁸⁴	NA
Cell fusion	24–48 h (<i>Nestin</i> , <i>Glur6</i> in NP–ES cell hybrids) ⁹²	24–48 h (Oct4–GFP reporter) ⁹²	About 9 d ⁹⁴	NA	G9a ⁹⁶ , Jhdmd2a ⁹⁶ , PRC1 and PRC2 (ref. 97)	Tet1 and Tet2 (ref. 93), Aid ²	Nanog ⁴⁵ , Oct4 (ref. 2), Sox2 not required ²	LIF, Wnt
PGC maturation <i>in vivo</i>	Around 2 d, between day 6.25–8.5 of gestation (<i>Hox</i> genes) ¹⁰⁰	1–2 d, between day 7.25–8.5 of gestation (<i>Sox2</i> , <i>Dppa3</i> , <i>Nanog</i>) ¹⁰⁰	About 3 d, between day 11.5–13.5 of gestation (X–GFP reporter) ¹⁰¹	H1 (ref. 99), H2A ^{Z99} , (Nap-1, Hira) ⁹⁹	Glp ¹⁰⁰ , Utx ¹⁴ , Prmt5 (ref. 103)	Dnmt3a, Dnmt3b and Dnmt3L ¹⁰⁷ ; Uhrf1 (ref. 107); Tet1 and Tet 2 (ref. 105); Nanog ¹⁰⁰ , Oct4 (ref. 100), Sox2 (ref. 100)	Blimp1 (ref. 100), Prdm14 (ref. 100), Tcfap2c ¹⁰⁰ , Nanog ¹⁰⁰ , Oct4 (ref. 100), Sox2 (ref. 100)	BMP ¹⁰⁰
PGC to EGC conversion <i>in vitro</i>	1–2 d (<i>Blimp1</i>) ¹⁰³	3–4 d (<i>Klf4</i> , <i>Stat3</i>) ¹⁰³	3–4 d (X–GFP reporter) ¹⁰¹	NA	NA	NA	NA	LIF or Stat3, FGF and SCF or ERK and GSK inhibition (2i) ⁷¹

EGC, embryonic germ cell; ES, embryonic stem; NA, not applicable; NP, neural progenitor; OKSM, Oct4, Klf4, Sox2 and c-Myc; PGC, primordial germ cell; SCNT, somatic cell nuclear transfer; TF, transcription factor. Note that some references refer to review articles covering the discussed molecule.

This finding suggests that ES cells, like oocytes, contain additional reprogramming factors that are limiting during iPSC formation. Corroborating this point, Nanog overexpression promotes cell fusion reprogramming and drives maturation of iPSCs, whereas its absence blunts both types of reprogramming⁴⁵. Surprisingly, Oct4 protein is required, whereas Sox2 protein is dispensable for fusion-mediated reprogramming, identifying a notable difference to induced pluripotency⁹⁵. It is also important to mention that the reprogramming of hybrids is not complete on activation of somatic pluripotency genes 2 days after fusion. Activation of the silenced X chromosome in female hybrids takes several days, which is reminiscent of the delayed X reactivation observed during *Xenopus* SCNT⁹⁶. Consistently, transcriptome-wide analysis of hybrid formation has documented that some silenced ES-cell-associated genes are activated more rapidly than others⁹⁷. These results are in line with the sequential activation of pluripotency-associated genes during iPSC derivation⁴⁵, and probably reflect different chromatin accessibility of associated genomic loci to ES-cell-derived reprogramming factors.

At the chromatin level, inhibition of the H3K9 HMT G9a or overexpression of the H3K9 HDM Jhdmd2a increases cell-fusion-directed reprogramming, which is in accordance with their supportive roles in iPSC generation⁹⁸. Conversely, depletion of PRC1 or PRC2 subunits decreases cell-fusion-mediated reprogramming and induced pluripotency, underscoring the general importance of H3K27me3-mediated gene repression for the acquisition of pluripotency⁹⁹. Collectively, these findings support the interpretation that iPSC formation and cell-fusion-directed reprogramming face similar epigenetic barriers and are stimulated by the same transcription factors, which is consistent with the fact that iPSC factors were initially identified in ES cells³. On the basis of accelerated kinetics of pluripotency activation in hybrids compared with nascent iPSCs, it should be feasible to devise genetic screens to identify other Nanog-like molecules that are limiting for efficient transcription-factor-induced reprogramming.

Primordial germ cell reprogramming

Primordial germ cell (PGC) maturation represents yet another type of reprogramming that occurs naturally and encompasses major epigenetic remodelling events that prepare the developing germ line for

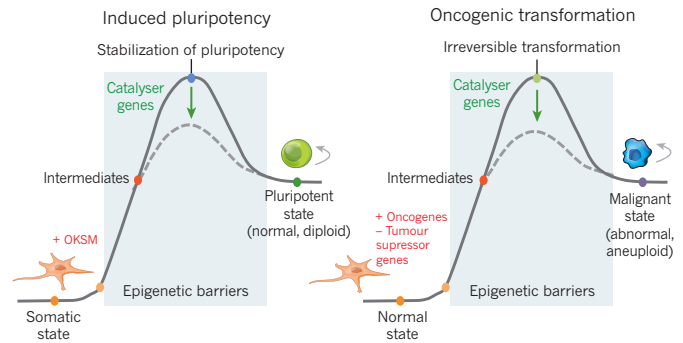
totipotency^{100,101}. Remarkably, PGCs that have completed reprogramming exhibit two active X chromosomes in females¹⁰², ES-cell-like transcriptional patterns and bivalent domains¹⁰³. This includes expression of potent reprogramming factors, such as Oct4, Sox2, Nanog and Prdm14 (refs 101, 104). Each of these factors is essential for PGC formation *in vivo*, although their precise roles in PGC reprogramming remain elusive. Importantly, PGCs are unipotent *in vivo* (they can only produce oocyte or sperm) but they have the unique potential to give rise to pluripotent stem cells, coined embryonic germ cells (EGCs), on explantation in culture¹. Expression of these ES-cell-associated transcription factors might thus endow PGCs with the latent ability to acquire pluripotency on isolation from the gonads and exposure to appropriate extracellular cues. Because germ cell reversion into EGCs rarely occurs *in vivo*, except in cases of spontaneous teratocarcinomas (pluripotent tumours), potent mechanisms must be in place to preserve the PGC state. Blimp1 is a possible candidate molecule, owing to its role as a repressor of c-Myc and Klf4 expression in PGCs¹⁰⁴. It should be possible to test this hypothesis by assessing whether acute loss of Blimp1 is sufficient to convert PGCs into EGCs *in vitro* and to cause teratocarcinomas *in vivo*. Notably, Blimp1's putative role in suppressing the acquisition of a pluripotent state in PGCs might be taken over by the transcription factor Dmrt1 at subsequent stages of male germline development. Male mice lacking this transcription factor in the germ line aberrantly express pluripotency factors and develop testicular teratomas with almost full penetrance¹⁰⁵. The notion that Blimp1 and Dmrt1 might actively resist acquisition of pluripotency is analogous to the inhibitory effect that somatic transcription factors have during iPSC formation¹⁹.

With respect to chromatin dynamics, global loss of H3K9 methylation is among the most striking changes in developing PGCs¹⁰⁰. Notably, downregulation of H3K9 HMTs seems to be crucial for efficient reprogramming in the context of PGC specification¹⁰¹, cell fusion⁹⁸ and iPSC formation^{24,36,37}. This particular chromatin alteration may therefore represent a general requirement for cellular reprogramming in these very different cellular contexts. Similarly, loss of inhibitory H3K27 methylation at pluripotency loci through catalysis by Utx seems to be another required step during both PGC reprogramming and iPSC formation¹⁴. Furthermore, pluripotency gene suppression by the NURD component

BOX 1

Induced pluripotency and malignant transformation

Here (Box Fig.) reprogramming of somatic cells to pluripotency and transformation of normal cells into malignant cells are illustrated as biochemical reactions with defined reactants (somatic cells and reprogramming factors or cancer genes) and products (iPSCs or cancer cells). Reprogramming is initiated by the overexpression of Oct4, Klf4, Sox2 and c-Myc (OKSM), whereas cellular transformation involves the activation of oncogenes and/or repression of tumour suppressor genes. Intermediates (coloured circles on grey lines) of each reaction remain largely elusive. Both processes need to overcome epigenetic barriers that stabilize the somatic state. Once a certain combination of epigenetic changes have been acquired (analogous to a rise in energy in a biochemical reaction), cells assume a stable new identity (iPSCs or cancer cells). The efficiency of this reaction can be modulated by manipulation of additional genes ('catalyser genes'). At a cellular level, both induced pluripotency and tumorigenesis are multi-step processes that require proliferation and result in a change of cell identity or differentiation potential. The 'end product' is in both cases an immortal cell with tumorigenic potential. However, cancer cells almost always acquire genetic aberrations and become aneuploid whereas iPSCs retain a normal diploid genome. At a molecular level, many cancer cells have, like iPSCs, reduced levels of H3K9 methylation and altered DNA methylation patterns compared



with differentiated cells. Overall, these findings suggest that nascent iPSCs and premalignant cells face some of the same epigenetic barriers to alter cell identity. This notion may explain why the same epigenetic regulators, such as Utx, macroH2A, Jhdmlb, Ezh2, Tet2 or Dnmts, are involved in both processes. This idea is also consistent with the finding that certain somatic progenitor and stem cells are more susceptible to tumorigenesis and reprogramming than differentiated cells, indicating a more permissive epigenetic environment (see 'Induced pluripotency and tumorigenesis' for details).

Mbd3 has recently been identified as a major reprogramming barrier during both EGC and iPSC induction⁵⁶. Last, Tet1 and Tet2 enzymes have been functionally associated with DNA demethylation in *in vitro*-derived PGCs¹⁰⁶, revealing similarities with SCNT⁸⁴ and cell fusion⁹⁴. However, it is noteworthy that genetic loss of both Tet1 and Tet2 is not essential for viability and fertility *in vivo*¹⁰⁷, suggesting compensation by other mechanisms. Indeed, passive demethylation was recently suggested to contribute to PGC reprogramming through downregulation of the *de novo* methyltransferases Dnmt3a and Dnmt3b and the Dnmt1 cofactor Uhrf1 (ref. 108). Combined with the enhancing effects of Tet1 and Tet2 overexpression⁴⁴, and Dnmt1 and Dnmt3a depletion^{31,42} on iPSC formation, these data show that the erasure of somatic DNA methylation patterns is a general roadblock for successful epigenetic reprogramming in different cellular settings. Cells use a combination of 'passive' and 'active' DNA demethylation strategies to overcome this barrier, although their relative contribution varies depending on the reprogramming context.

Induced pluripotency and tumorigenesis

Several lines of evidence support the idea that induced pluripotency and transformation are related processes at a cellular level (Box 1). Reprogramming, like cancer, is a rare, multi-step process that ultimately leads to the formation of a small population of immortal cells with tumorigenic potential; iPSCs, like ES cells, have the ability to give rise to teratomas (benign tumours containing derivatives of the three germ layers) on transplantation under the skin of mice¹. Another similarity between reprogramming and some types of cancer is the observation that somatic stem and progenitor cells are more susceptible to both iPSC formation^{109,110} and tumorigenesis^{111,112} compared with mature cells. This observation may indicate that the epigenetic state of the starting cell provides a permissive environment for both oncogenic and reprogramming factors. Furthermore, transcription-factor-mediated reprogramming induces a metabolic switch from an oxidative to a glycolytic state typical of most cancer cells¹¹³. Last, teratocarcinomas represent a special type of cancer that originates from transformed PGCs and contains pluripotent cells, documenting a rare example of spontaneous reprogramming of committed (germ) cells into pluripotent malignant

cells¹. An important distinction, however, between these examples of cancer and induced pluripotency is that iPSCs are normal diploid cells that support development when re-introduced into embryos, whereas most cancer cells are aneuploid and characterized by aberrant differentiation patterns. It should thus be informative to study those chromatin and epigenetic events that transiently endow iPSCs with immortality and tumorigenic properties in addition to increased differentiation potential, as this might lead to new strategies to reverse malignancy.

Molecular data support the cellular commonalities between reprogramming and malignancy. First, each of the four classic reprogramming factors has been shown to be oncogenic in mice, and some of these genes are amplified or mutated in human cancer¹¹⁴, suggesting that they might destabilize cell state in tumours akin to their role in reprogramming. Chromatin regulators that cooperate with OKSM during reprogramming, such as Jhdmlb and macroH2A, have also been associated with tumorigenesis. Jhdmlb expression has been causally linked with myeloid transformation of haematopoietic progenitors through silencing of the *ink4b* gene¹¹⁵ and with pancreatic adenocarcinoma formation through silencing of developmental genes in collaboration with PRC2 (ref. 116). In contrast to the cancer- and reprogramming-promoting role of Jhdmlb, expression of macroH2A provides a barrier for both iPSC formation and the malignant progression of melanoma cells. MacroH2A's promoting effect on melanoma invasion is, in part, exerted through upregulation of the cell cycle regulator CDK8, which differs from the pluripotency genes targeted by this histone variant during induced pluripotency¹¹⁷. Hence, these and several other examples¹¹⁴ document that premalignant cells and nascent iPSCs target some of the same chromatin regulators to manipulate cell identity, although their targets may vary.

Both cellular reprogramming and cancer are also characterized by similar global changes in chromatin structure and DNA methylation. Cancer cells, like ES cells, are devoid of LOCKs compared with normal differentiated cells²⁷. Given the observation that H3K9 methylation is a major barrier for iPSC formation^{24,36,37}, this finding suggests that many cancers have acquired a developmentally more primitive epigenetic state that might be required for the maintenance of malignancy. Another hallmark of most cancer genomes is altered methylation patterns, which

can manifest as aberrant hypermethylation or hypomethylation. In this regard, it is interesting to mention that reduced methylation levels, induced by hypomorphic expression of *Dnmt1*, cause T-cell lymphomas in mice¹¹⁸ and promote iPSC formation *in vitro*⁴². Likewise, mutations in DNMT3A have been observed in AML¹¹⁴, and knockdown of the gene encoding this enzyme facilitates human reprogramming into iPSCs³¹. These similarities between reprogramming and tumorigenesis are thus further consistent with the view that cancer cells need to override some of the same somatic barriers as iPSCs to alter cellular states.

Of note, the correlation between reprogramming and cancer is not absolute. In fact, some epigenetic regulators and histone modifications have been shown to have opposite roles during reprogramming and malignancy. For example, loss of Tet2 causes myeloid transformation in mice¹¹⁹, consistent with a tumour suppressor function, whereas depletion of Tet2 protein abrogates reprogramming⁴³. Similarly, the H3K79 methyltransferase Dot1L promotes leukaemia formation induced by MLL–AF9 translocations¹²⁰, although it prevents iPSC formation³¹. Components of the SWI–SNF complex, which facilitates reprogramming, act as potent tumour suppressors and are frequently mutated in cancer¹²¹, indicating opposite functions. Last, genome-wide methylation analyses of somatic cells and iPSCs have identified a set of reprogramming-specific differentially methylated regions (R-DMRs), which showed significant overlap with DMRs that changed during the transformation of normal into malignant cells¹²². However, R-DMRs that become hypomethylated and bivalent during iPSC generation are typically hypermethylated in cancer, whereas R-DMRs that become hypermethylated in iPSCs lack bivalent marks and are usually hypomethylated in cancer. Given the importance of bivalently marked genes in multi-lineage differentiation, their methylation silencing in cancer may be a secure way to keep cells in a self-perpetuating undifferentiated state, whereas this change would be detrimental for pluripotency in iPSCs. In summary, the discrepancies between tumorigenesis and reprogramming are probably explained by the strongly stage- and cell-context dependent roles that these enzymes and their modifications have during tumorigenesis and reprogramming.

A testable prediction based on the above-mentioned observations is that some cancer cells should be susceptible to epigenetic reprogramming into a non-malignant state. Indeed, both SCNT and iPSC experiments demonstrated that malignant cells, such as melanoma¹²³ and medulloblastoma¹²⁴, can be reprogrammed into a pluripotent state that supports differentiation into a number of normal cell types. Thus, some cancers are not irreversibly locked in a tumorigenic state but instead amenable to epigenetic reversion into a phenotypically normal state.

Outlook

Extensive functional genomics and screening approaches over the past few years have provided important insights into the epigenetic mechanisms that occur during normal, induced and pathological examples of cell fate change. Whereas the drivers of cell fate change may be quite different in distinct contexts (for example, OKSM in reprogramming, unidentified factors during SCNT and oncogenes in tumorigenesis), the resultant chromatin and epigenetic changes leading to altered cell identities are often conserved. This observation probably underlies the fact that different reprogramming approaches face some of the same molecular barriers that have been established during development and terminal differentiation to resist aberrant cell fate changes. We therefore conclude that transcription-factor-induced reprogramming provides a powerful tool to interrogate those chromatin and epigenetic mechanisms that stabilize cell fates during development and that become corrupted in cancer. These analyses have implications for both regenerative medicine and cancer biology. A better understanding of the molecular steps leading to pluripotency and the roadblocks resisting cell fate change in different contexts have already allowed researchers to interfere in a rationalized way with defined molecules or pathways to promote or prevent desired cell fate changes^{50–52,92}. An interesting challenge in the future will be to isolate and stabilize intermediate stages

of reprogramming, which might represent natural or artificial cellular states with increased differentiation potential. Dissecting the molecular roadblocks of reprogramming also has relevance for the study and treatment of cancer. Given that premalignant cells use some of the same epigenetic mechanisms as nascent iPSCs to change cell identity, their manipulation may lead to new strategies that reverse malignancy by altering cellular state rather than cell survival. Although the concept of reprogramming cancer cells to pluripotency has already been demonstrated, additional work is needed to develop more specific approaches that reverse malignant cells into a non-pluripotent state by targeting defined transcription factors or epigenetic regulators. Recent work on the conversion of leukaemia and lymphoma cells into non-tumorigenic, quiescent macrophages by a single transcription factor¹²⁵ is a promising step in this direction. ■

Received 3 July; accepted 5 August 2013.

1. Stadtfeld, M. & Hochedlinger, K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev.* **24**, 2239–2263 (2010).
 2. Yamanaka, S. & Blau, H. M. Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **465**, 704–712 (2010).
 3. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
 4. Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151–159 (2008).
 5. Stadtfeld, M., Maherali, N., Breault, D. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* **2**, 230–240 (2008).
 6. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55–70 (2007).
 7. Orkin, S. H. & Hochedlinger, K. Chromatin connections to pluripotency and cellular reprogramming. *Cell* **145**, 835–850 (2011).
 8. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnol.* **26**, 101–106 (2008).
 9. Wernig, M., Meissner, A., Cassady, J. & Jaenisch, R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* **2**, 10–12 (2008).
 10. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617–1632 (2012).
 11. Sridharan, R. *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136**, 364–377 (2009).
 12. Lin, C. Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56–67 (2012).
 13. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
 14. Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413 (2012).
 15. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–1222 (2012).
- This study performed the first single cell transcriptional analysis of nascent iPSCs and identified an early stochastic and a late deterministic phase.**
16. Shu, J. *et al.* Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**, 963–975 (2013).
- This report demonstrated that iPSCs can be derived through ectopic expression of lineage specifiers rather than conventional pluripotency factors.**
17. Loh, K. M. & Lim, B. A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell* **8**, 363–369 (2011).
 18. Hanna, J. *et al.* Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**, 250–264 (2008).
 19. Hikichi, T. *et al.* Transcription factors interfering with dedifferentiation induce cell type-specific transcriptional profiles. *Proc. Natl Acad. Sci. USA* **110**, 6412–6417 (2013).
 20. Gifford, C. A. *et al.* Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**, 1149–1163 (2013).
 21. Xie, W. *et al.* Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **153**, 1134–1148 (2013).
- References 20 and 21 established comprehensive chromatin and DNA methylation maps of pluripotent and diverse lineage-restricted cells.**
22. Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642–654 (2013).
 23. Koche, R. P. *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* **8**, 96–105 (2011).
 24. Soufi, A., Donahue, G. & Zaret, K. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
- This publication determined how somatic chromatin structure influences access by reprogramming factors and identified important barriers to iPSC formation.**

25. Polo, J. M. & Hochedlinger, K. When fibroblasts MET iPSCs. *Cell Stem Cell* **7**, 5–6 (2010).
26. Taberlay, P. C. *et al.* Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* **147**, 1283–1294 (2011).
27. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genet.* **41**, 246–250 (2009).
28. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
29. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).
30. Jia, J. *et al.* Regulation of pluripotency and self-renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell* **151**, 576–589 (2012).
31. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
32. Fragola, G. *et al.* Cell reprogramming requires silencing of a core subset of polycomb targets. *PLoS Genet.* **9**, e1003292 (2013).
33. Wang, T. *et al.* The histone demethylases Jhdmla/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* **9**, 575–587 (2011).
34. Zhao, W. *et al.* Jmjd3 inhibits reprogramming by upregulating expression of INK4a/Arf and targeting PHF20 for ubiquitination. *Cell* **152**, 1037–1050 (2013).
35. Liang, G., He, J. & Zhang, Y. Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nature Cell Biol.* **14**, 457–466 (2012).
36. Chen, J. *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genet.* **45**, 34–42 (2013).
37. Sridharan, R. *et al.* Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1 γ in reprogramming to pluripotency. *Nature Cell Biol.* **15**, 872–882 (2013).
- Reports 36 and 37 identified H3K9 methylation as a major chromatin barrier of cellular reprogramming into iPSCs.**
38. Ang, Y. S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).
39. Esch, D. *et al.* A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nature Cell Biol.* **15**, 295–301 (2013).
40. Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013).
41. Pawlak, M. & Jaenisch, R. *De novo* DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev.* **25**, 1035–1040 (2011).
42. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
43. Doege, C. A. *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**, 652–655 (2012).
44. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).
45. Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
46. Gao, Y. *et al.* Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
47. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
48. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
49. Skene, P. J. & Henikoff, S. Histone variants in pluripotency and disease. *Development* **140**, 2513–2524 (2013).
50. Gaspar-Maia, A. *et al.* MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nature Commun.* **4**, 1565 (2013).
51. Pasque, V. *et al.* Histone variant macroH2A marks embryonic differentiation *in vivo* and acts as an epigenetic barrier to induced pluripotency. *J. Cell Sci.* **125**, 6094–6104 (2012).
52. Barrero, M. J. *et al.* Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. *Cell Rep.* **3**, 1005–1011 (2013).
53. Singhal, N. *et al.* Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* **141**, 943–955 (2010).
54. Gaspar-Maia, A. *et al.* Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* **460**, 863–868 (2009).
55. Luo, M. *et al.* NuRD Blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells* **31**, 1278–1286 (2013).
56. Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* <http://dx.doi.org/10.1038/nature12587> (18 September 2013).
- This study reports that the depletion of the NURD component Mbd3 during iPSC induction is sufficient to render reprogramming a deterministic process.**
57. Yildirim, O. *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* **147**, 1498–1510 (2011).
58. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
59. Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
60. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
61. Apostolou, E. *et al.* Genome-wide chromatin interactions of the *Nanog* locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* **12**, 699–712 (2013).
62. Wei, Z. *et al.* Klf4 organizes long-range chromosomal interactions with the *Oct4* locus in reprogramming and pluripotency. *Cell Stem Cell* **13**, 36–47 (2013).
63. Zhang, H. *et al.* Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell* **13**, 30–35 (2013).
- References 61–63 identified pluripotency-specific, genome-wide interaction networks around pluripotency loci and implicate mediator and cohesin in their regulation.**
64. Levasseur, D. N., Wang, J., Dorschner, M., Stamatoyannopoulos, J. & Orkin, S. Oct4 dependence of chromatin structure within the extended *Nanog* locus in ES cells. *Genes Dev.* **22**, 575–580 (2008).
65. Schoenfelder, S. *et al.* Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nature Genet.* **42**, 53–61 (2010).
66. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
67. Silva, J. *et al.* Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* **6**, e253 (2008).
68. Yang, J. *et al.* Stat3 activation is limiting for reprogramming to ground state pluripotency. *Cell Stem Cell* **7**, 319–328 (2010).
69. Ichida, J. K. *et al.* A small-molecule inhibitor of Tgf- β signaling replaces *Sox2* in reprogramming by inducing *Nanog*. *Cell Stem Cell* **5**, 491–503 (2009).
70. Maherali, N. & Hochedlinger, K. Tgf β signal inhibition cooperates in the induction of iPSCs and replaces *Sox2* and *cMyc*. *Curr. Biol.* **19**, 1718–1723 (2009).
71. Ho, L. *et al.* esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nature Cell Biol.* **13**, 903–913 (2011).
72. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nature Struct. Mol. Biol.* **20**, 311–316 (2013).
73. Yamaji, M. *et al.* PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368–382 (2013).
- References 71 and 72 characterize the effect of combined Gsk3/Mapk signal inhibition ('2i') on the epigenome of ES cells and provide an intriguing link to Prdm14.**
74. Ho, R., Papp, B., Hoffman, J. A., Merrill, B. J. & Plath, K. Stage-specific regulation of reprogramming to induced pluripotent stem cells by wnt signaling and T cell factor proteins. *Cell Rep.* **3**, 2113–2126 (2013).
75. Esteban, M. A. *et al.* Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**, 71–79 (2010).
76. Stadtfeld, M. *et al.* Ascorbic acid prevents loss of Dlk1–Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nature Genet.* **44**, 398–405 (2012).
77. Blaschke, K. *et al.* Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226 (2013).
78. O'Malley, J. *et al.* High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* **499**, 88–91 (2013).
79. Hansson, J. *et al.* Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep.* **2**, 1579–1592 (2012).
80. Tanabe, K., Nakamura, M., Narita, M., Takahashi, K. & Yamanaka, S. Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc. Natl Acad. Sci. USA* **110**, 12172–12179 (2013).
81. Golipour, A. *et al.* A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* **11**, 769–782 (2012).
82. Hou, P. *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651–654 (2013).
- This provocative report suggests that induced pluripotency can be achieved with four small compounds and without the use of transcription factors.**
83. Egli, D. *et al.* Reprogramming within hours following nuclear transfer into mouse but not human zygotes. *Nature Commun.* **2**, 488 (2011).
84. Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
85. Wu, G. *et al.* Establishment of totipotency does not depend on Oct4A. *Nature Cell Biol.* **15**, 1089–1097 (2013).
- The authors of this study discovered that Oct4 is dispensable for reprogramming somatic nuclei by SCNT, which pointed to important mechanistic differences with induced pluripotency.**
86. Maekawa, M. *et al.* Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–229 (2011).
87. Jullien, J., Astrand, C., Halley-Stott, R. P., Garrett, N. & Gurdon, J. B. Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. *Proc. Natl Acad. Sci. USA* **107**, 5483–5488 (2010).
88. Lu, X. *et al.* *Drosophila* H1 regulates the genetic activity of heterochromatin by recruitment of Su(var)3–9. *Science* **340**, 78–81 (2013).

89. Yang, S. M., Kim, B. J., Norwood Toro, L. & Skoultschi, A. I. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl Acad. Sci. USA* **110**, 1708–1713 (2013).
90. Nashun, B., Akiyama, T., Suzuki, M. G. & Aoki, F. Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos. *Epigenetics* **6**, 1489–1497 (2011).
91. Jullien, J. et al. HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to *Xenopus* oocytes. *Epigenetics Chromatin* **5**, 17 (2012).
92. Pasque, V., Gillich, A., Garrett, N. & Gurdon, J. B. Histone variant macroH2A confers resistance to nuclear reprogramming. *EMBO J.* **30**, 2373–2387 (2011).
93. Han, D. W. et al. Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. *Stem Cells* **26**, 445–454 (2008).
94. Piccolo, F. M. et al. Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell* **49**, 1023–1033 (2013).
95. Pereira, C. F. et al. Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS Genet.* **4**, e1000170 (2008).
96. Do, J. T. et al. Enhanced reprogramming of Xist by induced upregulation of Tsix and Dnmt3a. *Stem Cells* **26**, 2821–2831 (2008).
97. Foshay, K. M. et al. Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Mol. Cell* **46**, 159–170 (2012).
98. Ma, D. K., Chiang, C. H., Ponnusamy, K., Ming, G. L. & Song, H. G9a and Jhdmd2a regulate embryonic stem cell fusion-induced reprogramming of adult neural stem cells. *Stem Cells* **26**, 2131–2141 (2008).
99. Pereira, C. F. et al. ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* **6**, 547–556 (2010).
100. Hajkova, P. et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877–881 (2008).
101. Magnúsdóttir, E., Gillich, A., Grabole, N. & Surani, M. A. Combinatorial control of cell fate and reprogramming in the mammalian germline. *Curr. Opin. Genet. Dev.* **22**, 466–474 (2012).
102. Chuva de Sousa Lopes, S. M. et al. X chromosome activity in mouse XX primordial germ cells. *PLoS Genet.* **4**, e30 (2008).
103. Sachs, M. et al. Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline *in vivo*. *Cell Rep.* **3**, 1777–1784 (2013).
104. Durcova-Hills, G., Tang, F., Doody, G., Tooze, R. & Surani, M. A. Reprogramming primordial germ cells into pluripotent stem cells. *PLoS ONE* **3**, e3531 (2008).
105. Krentz, A. D. et al. The DM domain protein DMRT1 is a dose-sensitive regulator of fetal germ cell proliferation and pluripotency. *Proc. Natl Acad. Sci. USA* **106**, 22323–22328 (2009).
106. Hackett, J. A. et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
107. Dawlaty, M. M. et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–323 (2013).
108. Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. & Saitou, M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**, 340–353 (2013).
109. Eminli, S. et al. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nature Genet.* **41**, 968–976 (2009).
110. Tan, K. Y., Eminli, S., Hettmer, S., Hochedlinger, K. & Wagers, A. J. Efficient generation of iPS cells from skeletal muscle stem cells. *PLoS ONE* **6**, e26406 (2011).
111. White, A. C. et al. Defining the origins of Ras/p53-mediated squamous cell carcinoma. *Proc. Natl Acad. Sci. USA* **108**, 7425–7430 (2011).
112. Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).
113. Folmes, C. D. et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab.* **14**, 264–271 (2011).
114. Suvà, M. L., Riggi, N. & Bernstein, B. E. Epigenetic reprogramming in cancer. *Science* **339**, 1567–1570 (2013).
115. He, J., Nguyen, A. T. & Zhang, Y. KDM2b/JHDM1b, an H3K36me2-specific demethylase, is required for initiation and maintenance of acute myeloid leukemia. *Blood* **117**, 3869–3880 (2011).
116. Tzatsos, A. et al. KDM2B promotes pancreatic cancer via Polycomb-dependent and -independent transcriptional programs. *J. Clin. Invest.* **123**, 727–739 (2013).
117. Kapoor, A. et al. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. *Nature* **468**, 1105–1109 (2010).
118. Gaudet, F. et al. Induction of tumors in mice by genomic hypomethylation. *Science* **300**, 489–492 (2003).
119. Moran-Crusio, K. et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
120. Krivtsov, A. V. et al. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* **14**, 355–368 (2008).
121. Kadoch, C. et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genet.* **45**, 592–601 (2013).
122. Doi, A. et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genet.* **41**, 1350–1353 (2009).
123. Hochedlinger, K. et al. Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev.* **18**, 1875–1885 (2004).
124. Stricker, S. H. et al. Widespread resetting of DNA methylation in glioblastoma-initiating cells suppresses malignant cellular behavior in a lineage-dependent manner. *Genes Dev.* **27**, 654–669 (2013).
125. Rapino, F. et al. C/EBP α induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep.* **3**, 1153–1163 (2013).

Acknowledgements We thank A. Bortvin, D. Egli, V. Pasque, B. Payer, B. Ren, X. Wei, M. Tahiliani and members of the laboratory for helpful suggestions, and B. Lowry, K. Plath and M. Stadtfeld for critical reading of the manuscript. We apologize to those colleagues whose work we could not cite due to space constraints. Support to E.A. was by HHMI and to K.H. by HHMI and NIH (R01HD058013).

Author Information Reprints and permission information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/vfkh3f. Correspondence should be addressed to K.H. (khochedlinger@helix.mgh.harvard.edu).

TET enzymes, TDG and the dynamics of DNA demethylation

Rahul M. Kohli^{1,2} & Yi Zhang^{3,4,5,6,7}

DNA methylation has a profound impact on genome stability, transcription and development. Although enzymes that catalyse DNA methylation have been well characterized, those that are involved in methyl group removal have remained elusive, until recently. The transformative discovery that ten-eleven translocation (TET) family enzymes can oxidize 5-methylcytosine has greatly advanced our understanding of DNA demethylation. 5-Hydroxymethylcytosine is a key nexus in demethylation that can either be passively depleted through DNA replication or actively reverted to cytosine through iterative oxidation and thymine DNA glycosylase (TDG)-mediated base excision repair. Methylation, oxidation and repair now offer a model for a complete cycle of dynamic cytosine modification, with mounting evidence for its significance in the biological processes known to involve active demethylation.

Two competing demands on the genome are the need for stability and for flexibility. Because the functional programs of cells are encoded by the genome, this information must be faithfully propagated both through development and across generations. At the same time, the genome of a multicellular organism must encode for diverse cell types, each of which must be capable of responding to a changing environment. These latter functions require the capacity for adaptive regulation of gene expression, which can be achieved by the transcription factor complexes that bind DNA, by the packaging of DNA into chromatin and by dynamic covalent modifications to either histones or DNA itself.

Covalent modification of DNA, in particular, helps to provide a means for functional variability while maintaining the information content of the base. One of the best-studied covalent modifications on DNA is 5-methylcytosine (5mC), a mark deposited by DNA methyltransferase (DNMT) enzymes¹. In mammalian genomes, 5mC exists mostly in the CpG dinucleotide context and about 70–80% of CpGs are methylated. DNMTs can both introduce methylation marks (*de novo* methylation) and maintain them after the genome is replicated (maintenance methylation), making DNA methylation a long-term and potentially heritable mark¹. Conventionally, 5mC is associated with a transcriptionally repressed chromatin state, and DNA methylation at specific genomic loci, including lineage-specific genes, can help to shape a cellular program during development². 5mC-mediated long-term gene silencing also contributes to genomic imprinting, X-chromosome inactivation and suppression of mobile genetic elements^{3,4}.

DNA methylation is relatively stable compared with most histone modifications. Nevertheless, loss of DNA methylation, or DNA demethylation, has been observed in different biological contexts and this alteration can take place actively or passively. Active DNA demethylation refers to an enzymatic process that removes or modifies the methyl group from 5mC. By contrast, passive DNA demethylation refers to loss of 5mC during successive rounds of replication in the absence of functional DNA methylation maintenance machinery. Although passive DNA demethylation is generally understood and accepted, the evidence for active DNA demethylation and how it occurs has been controversial^{5,6}. In part, this controversy has been due to the cacophony of enzymes and pathways implicated in demethylation.

However, a series of recent discoveries has begun to harmonize, and thereby greatly advance, our understanding of active DNA demethylation. Here, we review these significant discoveries, their biological implications and the promising areas for further exploration.

DNA demethylation and historical mechanisms

Several reviews have described the biological context in which active DNA demethylation may take place^{5–7}. Establishing and editing genomic methylation patterns seems to be particularly relevant in several stages of mammalian embryogenesis. Initially, after the sperm penetrates the egg and before the merging of paternal and maternal genomes, the paternal genome goes through a complex remodelling process that includes deposition of histone H3.3 and remodelling of DNA methylation patterns⁸. Here, a rapid loss of 5mC staining is observed in the paternal, but not the maternal, genome, suggesting an active 5mC editing process^{9,10}. After implantation, and early in development, a subset of posterior epiblast cells is instructed to become primordial germ cells (PGCs). PGCs have to go through a complex epigenetic reprogramming process, including erasure of genome-wide DNA methylation patterns¹¹, to prepare them for germ-cell-specific processes, such as meiosis. Besides global loss of DNA methylation in zygotes and PGCs, DNA demethylation has also been observed at specific loci in rapid response to environmental stimuli or in post-mitotic cells, supporting the relevance of active demethylation in various biological settings in the absence of cellular replication^{12–14}.

Many candidates from the known repertoire of DNA modifying enzymes have historically been proposed to function in DNA demethylation (see refs 5, 6, 15 and 16 for reviews). As we will discuss in the context of more recent discoveries, DNA cytosine deaminases that can introduce genomic mismatches, DNA glycosylases that can excise bases, other DNA repair factors and even DNMTs themselves have been suggested to be involved in DNA demethylation. Although there is some evidence to support a role for many of these DNA modifying pathways, these roles have often seemed specific to the individual biological system being examined. The lack of a unifying mechanistic process has led to ongoing dispute over the relative importance of these various pathways in DNA demethylation. Although these

¹Department of Medicine, Raymond and Ruth Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA. ²Department of Biochemistry and Biophysics, Raymond and Ruth Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA. ³Howard Hughes Medical Institute, WAB-1496, 200 Longwood Ave., Boston, Massachusetts 02115, USA. ⁴Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, Massachusetts 02115, USA. ⁵Division of Hematology/Oncology, Department of Pediatrics, Boston Children's Hospital, Boston, Massachusetts 02115, USA. ⁶Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁷Harvard Stem Cell Institute, 200 Longwood Ave., Boston, Massachusetts 02115, USA.

multiple candidate pathways remain areas of active exploration^{5,6,15,16}, in this Review we focus on recent developments that have brought new clarity to the field of DNA demethylation by elucidating pathways of oxidation-mediated demethylation.

TET-mediated oxidation of 5mC

The discovery of a family of enzymes that can modify 5mC through oxidation was a watershed moment in advancing our understanding of DNA demethylation mechanisms, introducing 5-hydroxymethylcytosine (5hmC) as a key intermediate in active demethylation pathways^{17–19}. This discovery was motivated by the study of two pathways involving oxidative modifications of T bases: one involving oxidative modifications to DNA, the other, demethylation of a nucleobase. In the parasite responsible for African sleeping sickness, *Trypanosoma brucei*, glucosylated 5-hydroxymethyluracil (Base J) functions in transcriptional regulation to modulate surface glycoprotein expression and thereby promote immune escape²⁰. Base J biosynthesis involves oxidation of T within DNA to 5-hydroxymethyluracil (5hmU) by JBP1 and JBP2, members of the Fe(II)/ α -ketoglutarate (α -KG)-dependent oxygenase family of enzymes. A second member of this oxygenase family, thymidine hydroxylase, acts instead on free T bases in a pyrimidine salvage pathway. Interestingly, the initial oxidation product, 5hmU, is subsequently further oxidized to 5-formyluracil and 5-carboxyluracil²¹. Decarboxylation by isoorotate decarboxylase²² completes a cycle from T to U with potential mechanistic parallels to 5mC demethylation⁵.

Bioinformatic analyses by Rao and colleagues^{17,23} revealed several mammalian paralogues of JBP1 and JBP2 (Fig. 1a). These enzymes belong to the TET family, which has previously been implicated in haematopoietic malignancies²⁴. Surprisingly, overexpression of TET1 was associated with a reduction in genomic 5mC, suggesting that, unlike its paralogues, TET1 recognized modified C, rather than T, bases in DNA¹⁷. Indeed, purified TET enzymes modified oligonucleotide substrates containing 5mC through oxidation, and the product was authenticated as 5hmC^{17,18} (Fig. 1b).

Although 5hmC had previously been observed in mammalian genomes²⁵, these earlier observations did not receive attention until the discovery of TET enzymes, because these catalysts are capable of purposefully generating this oxidized base. Moreover, 5hmC was shown to be readily detectable in mouse embryonic stem (ES) cells in a manner dependent on expression of TET^{17,18}. Even more strikingly, 5hmC was shown to be nearly 40% as abundant as 5mC in post-mitotic neuronal Purkinje cells¹⁹. Even though neuronal cells seem particularly enriched, accurate quantification methods have since demonstrated that 5hmC accumulates in most cell types, raising the possibility that this 'sixth base' in the genome may have a distinctive epigenetic role^{26,27}.

Reflecting on the discovery of TET, two groups noted that the initial assay conditions used for detecting 5hmC were not permissive for detecting further oxidative modifications^{28,29}. With alternative chromatographic conditions, additional products emerged, demonstrating that, like thymidine hydroxylase, TET was capable of iterative oxidation, yielding 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC)^{28,29}. Highly sensitive mass spectrometry established that these base modifications are detectable either with TET overexpression or, of more physiological relevance, within ES cells, although their levels are at least an order of magnitude less than those of 5hmC^{28,29}. Notably, 5hmC, 5fC and 5caC are chemically distinct modifications of C that could be specifically recognized by different DNA-binding proteins. The oxidized 5-substituents also have different steric and electronic properties, which can promote alternative nucleobase tautomers or, in the case of 5fC and 5caC, destabilize the *N*-glycosidic bond^{30,31}.

The TET family members (TET1, 2 and 3) each harbour a core catalytic domain (Fig. 1a), with a double-stranded β -helix fold that contains the crucial metal-binding residues found in the family of Fe(II)/ α -KG-dependent oxygenases³². In the putative mechanism based on the precedent of other family members (Fig. 1b), TET uses molecular oxygen as a substrate to catalyse oxidative decarboxylation of α -KG, thereby

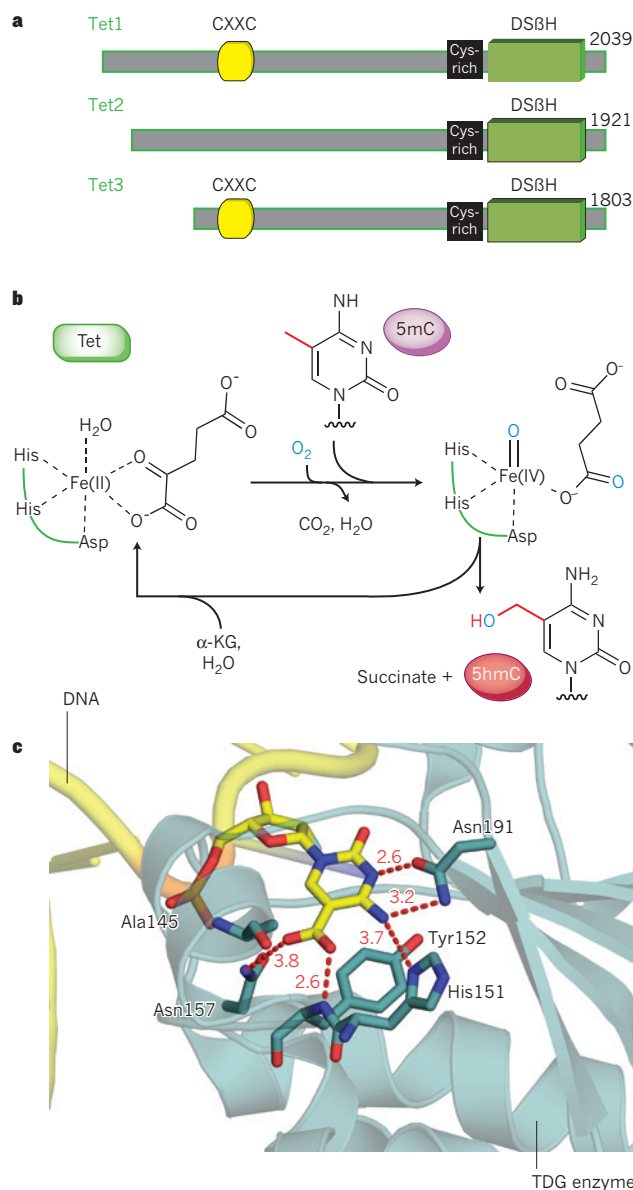


Figure 1 | TET and TDG function in oxidation and excision of modified C bases. **a**, Schematic of mouse Tet enzymes, showing the double-stranded β -helix fold core oxygenase domain, a preceding cysteine (Cys)-rich domain and a CXXC domain in Tet1 and Tet3. **b**, Catalytic pathway for generation of 5hmC by Tet enzymes. An active site Fe(II) (left) is bound by conserved His–His–Asp residues in Tet and coordinates water and α -ketoglutarate (α -KG). A two-electron oxidation of α -KG by molecular oxygen yields CO_2 and enzyme-bound succinate, and results in a high-valent Fe(IV)-oxo intermediate (right). The intermediate reacts with 5mC to yield 5hmC, with a net oxidative transfer of the single oxygen atom to the substrate, resulting in regeneration of the Fe(II) species. **c**, TDG specifically accommodates oxidized C bases. Shown is the active site of TDG, bound to DNA, containing a substrate analogue of 5caC (PDB 3UOB). Critical residues of the enzyme are labelled. The 5caC analogue is highlighted in yellow. Heteroatoms are shown with nitrogen (blue), oxygen (red) and phosphorus (orange). The distance of hydrogen bonds (dashed red lines) are measured in Å. In addition to several interactions with the Watson–Crick face of the base from Asn 191 and His 151, the carboxylate substituent in the 5-position is well-accommodated by the active site with a binding pocket defined by Ala 145, and hydrogen bonds from Asn 157 and the backbone amide of Tyr 152.

generating a reactive high-valent enzyme-bound Fe(IV)-oxo intermediate that converts 5mC to 5hmC. The core catalytic domain constitutes only a fraction of the large TET enzymes, suggesting the possibility that the non-catalytic domains may have regulatory functions (Fig. 1a). In all

TET isoforms, a cysteine-rich domain precedes the core and seems to be required for activity^{17,23}. TET1 and TET3 also contain a chromatin-associated CXXC domain that is known to bind CpG sequences, whereas TET2 partners with IDAX, an independent CXXC-containing protein^{33–35}.

Replication, repair and demethylation

Detection of oxidized 5mC bases within ES cells has suggested potential functional relevance for these bases in the dynamic regulation of the genome and led to the next question: how might these oxidized bases be altered to regenerate unmodified C? Three potential pathways for demethylation following 5mC oxidation have been entertained: passive dilution of the oxidized base, direct removal of the oxidized 5-position substituent and DNA repair-mediated excision of modified nucleotides.

As had been previously entertained with 5mC, passive dilution of 5hmC or the highly oxidized bases may contribute to demethylation. Indeed, significant evidence (discussed later) has pointed to the importance of this DNA replication-dependent pathway. Although confusion exists in the literature as to whether this pathway should be designated as active or passive (see Perspectives), we find it most useful to consider this as an active demethylation pathway that results from active modification of 5mC followed by passive dilution of the oxidized base to regenerate unmodified C in a replication-dependent manner.

What about pathways that might promote active restoration of unmodified C? Whereas direct removal of a methyl group has a high energetic barrier, the removal of the oxidized methyl group is more feasible. For example, similar to the precedent of isoorotate decarboxylase, decarboxylation of 5caC could revert the base to unmodified C. One study with isotopic labelling of 5caC has suggested this possibility³⁶; however, a 5caC decarboxylase has yet to be identified. Interestingly, in the absence of the methyl-donor S-adenosylmethionine (SAM), DNMTs can potentially promote the addition or removal of oxidized 5-position substituents,

including reacting with 5hmC *in vitro*³⁷. Thus, DNMTs could theoretically function in demethylation, raising interesting regulatory implications. The biological relevance of this 'reverse' DNMT reaction remains unknown because SAM is a general methyl-group donor that is present in all cell types.

On firmer ground, an alternative pathway for active restoration of C could involve DNA repair enzymes. Although pathways involving nucleotide excision repair have been considered in demethylation^{38,39}, the bulk of the focus has been on base excision repair (BER), a pathway involving the removal of an entire modified base and its subsequent repair to replace the residue with unmodified C (see ref. 40 for review of BER). Several key components of the BER pathway are present at crucial transitions of DNA methylation patterns^{41–43}, and this line of inquiry, as detailed in the next section, has proven fruitful.

TDG-mediated repair completes the cycle

The suggested involvement of BER in demethylation prompted an active search for glycosylase enzymes that might excise modified C bases. Plants use such a mechanism to excise 5mC directly, and some early reports suggested that either methyl-binding domain protein 4 (MBD4) or thymine DNA glycosylase (TDG) may have similar activity in mammals⁴⁴. Although these possibilities have since been discounted given these enzymes' marginal 5mC glycosylase activity, there is mounting biochemical evidence for a role of TDG in DNA demethylation⁴⁵. In particular, TDG has been shown to interact with numerous transcription factors, chromatin modifying enzymes and DNMTs, raising the possibility of a functional role for TDG in modulating gene transcription, either through its glycosylase activity or as a transcriptional coactivator⁴⁵.

After the discovery of TET, the next significant milestone in DNA demethylation came when two groups demonstrated that, unlike other

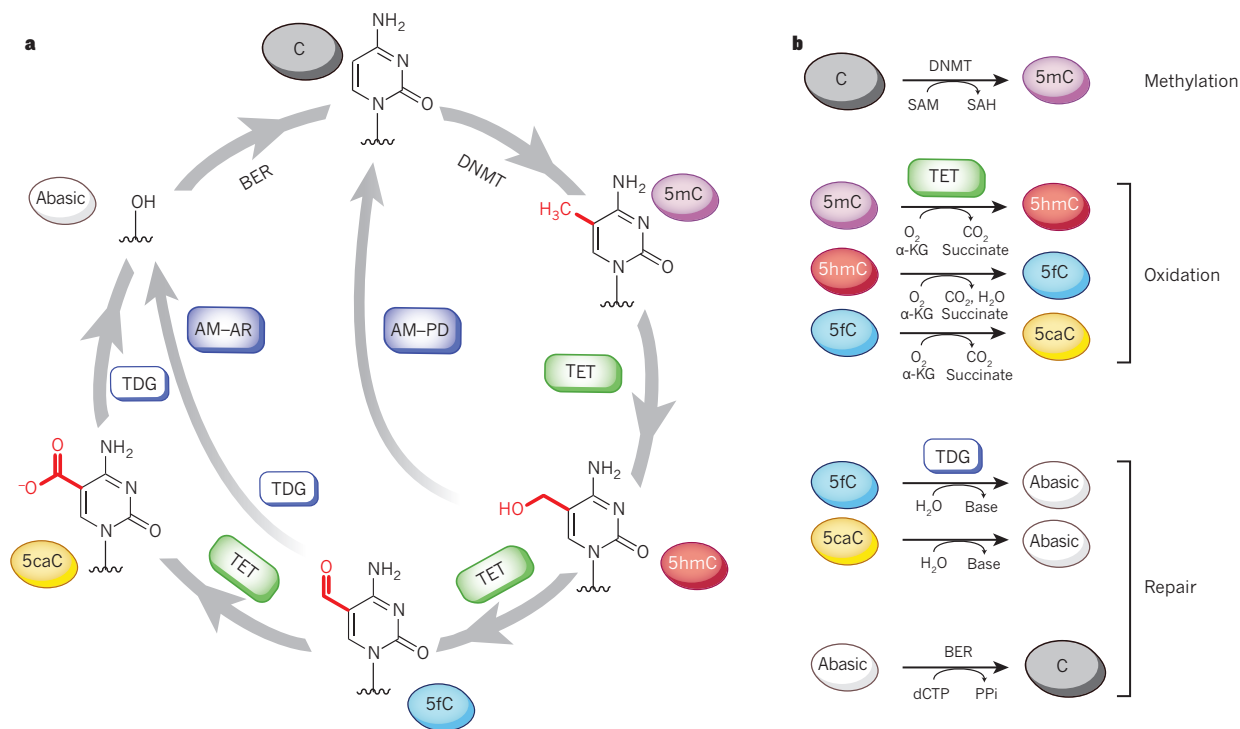


Figure 2 | A complete pathway for dynamic modifications of C. **a**, A biochemically validated pathway for modification of C within DNA is shown. 5mC bases, introduced by DNA methyltransferase (DNMT) enzymes, can be oxidized iteratively to 5hmC, 5fC and 5caC. In the pathway of active modification (AM) followed by passive dilution (PD), 5hmC is diluted in a replication-dependent manner to regenerate unmodified C. For clarity, PD of highly oxidized 5fC and 5caC is not depicted. In the pathway of AM followed by active restoration (AR), 5fC

or 5caC is excised by TDG generating an abasic site as part of the base excision repair (BER) process that regenerates unmodified C. **b**, The individual reactions in the pathway are shown with all reactants depicted. The BER pathway involves excision of the abasic site, replacement of the nucleotide using unmodified deoxycytidine triphosphate (dCTP) by a DNA polymerase (generating pyrophosphate, PPi) and ligation to repair the nick. α -KG, α -ketoglutarate; SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine.

DNA glycosylases, TDG is required for embryonic development^{46,47}. Molecular studies on TDG-null embryos, or a catalytically inactive mutant, have pointed to an epigenetic abnormality. Among other alterations, the mutants showed marked decreases in the expression of developmental transcription factors, such as *hox* gene family members, with perturbed methylation at their regulatory sequences^{46,47}. Although these genetic studies raised the possibility of TDG as an important player in DNA demethylation, the nature of its genomic target remained unclear.

TDG has long been the focus of compelling biochemical and structural studies because of its interesting role as a DNA repair enzyme that can remove a normal base, T, from a genomic T·G mismatch⁴⁸. Initial speculation therefore focused on the possibility that TDG activity would be coupled to deamination of 5mC or 5hmC because T·G or hmU·G mismatches were known substrates of TDG for which repair could regenerate unmodified C^{47,49}. In this model, AID/APOBEC enzymes, the adaptive or innate immune system enzymes that normally target unmodified C, were considered to be the likely candidates for catalysing deamination. Indeed, some studies have suggested a role for deaminases in the reprogramming of stem cells or in embryogenesis^{50–53}. These deamination-mediated pathways for demethylation could also involve the DNA damage response protein GADD45 or even MBD4 as an alternative glycosylase for excision of T·G mismatches^{38,53,54}, although evidence to the contrary exists^{55,56}. Notably, however, deamination of 5hmC by AID/APOBEC enzymes is not detectable *in vitro* or in cells^{57,58}, challenging the plausibility of proposed pathways that have invoked 5hmC deamination in DNA demethylation⁴⁹. By contrast, deamination of 5mC by AID/APOBEC family members does occur at a detectable rate *in vitro*, about 10-fold slower than with unmodified C⁵⁷. Although feasible, we anticipate that the role of deaminases in demethylation is probably limited given that 5mC is much less abundant than the unmodified C in mammalian genomes, as well as the enzyme's selectivity for single-stranded DNA and its preference for particular sequence contexts. This view is supported by the observation that there are no significant developmental defects associated with AID/APOBEC deficiency¹⁶.

A role for TDG in processing T·G mismatches potentially generated by AID/APOBEC deamination of 5mC remained one possible explanation for its requirement in embryonic development. However, the scope of TDG's role in demethylation was reconsidered on revisiting a previous observation that some correctly base paired, modified C bases can also be targeted by TDG⁵⁹. Specifically, C bases with 5-position substituents that destabilize the N-glycosidic bond by electronic effects, such as 5-fluorocytosine, have been shown to be efficiently excised by the glycosylase. These observations opened up the possibility that TDG could directly excise TET oxidation products. Indeed, although no significant *in vitro* base excision activity has been observed with C, 5mC and 5hmC, TDG has robust *in vitro* base excision activity on 5fC and 5caC properly base paired to G in duplex DNA^{29,60} (Fig. 1c). This *in vitro* activity is relevant in cells, as knockdown of the gene encoding TDG leads to elevated 5caC levels in ES cells²⁹. Furthermore, simultaneous TET and TDG overexpression in the HEK293 cell line leads to a depletion of TET-associated 5caC^{29,57} or 5fC⁵⁷. Thus, in a striking example of synergy, studies demonstrating a requirement for TDG in development could be reconciled with insights into TET-mediated oxidation. TDG, acting on TET-generated 5fC and 5caC, mediates the first biologically and biochemically validated, complete pathway for active DNA demethylation (Fig. 2).

Biochemical and biophysical studies have started to shed light on the molecular basis for excision of 5fC and 5caC by TDG. In line with earlier studies on TDG's requirements for excision⁵⁹, computational studies have suggested that 5fC and 5caC have destabilized N-glycosidic bonds relative to C, 5mC and 5hmC^{30,59}. TDG also seems to have structural features that mediate recognition of these oxidized C bases, including a binding pocket that can specifically accommodate the 5-carboxyl substituent⁶¹ (Fig. 1c). Interestingly, the determinants for

5fC excision seem to be separable from 5caC recognition³¹, an insight that will probably prove useful in assessing the relative importance of 5fC compared with 5caC to demethylation.

Revisiting biological roles for demethylation

As our biochemical knowledge of TET, TDG and other DNA modifying pathways has evolved, the many biological processes in which DNA demethylation seems to be relevant have been re-examined with a fresh perspective. 5hmC resides at a potentially crucial branch point in the DNA demethylation pathway (Box 1). Here we summarize recent advances, focusing on studies that help to establish the role of 5hmC in various biological and pathological settings in which dynamic DNA methylation takes place both globally and locally.

Pre-implantation global methylation dynamics

The specific and rapid loss of 5mC from the paternal genome of zygotes has been re-examined in light of the discovery of 5hmC (Fig. 3a). Immunostaining using a 5hmC-specific antibody revealed that loss of 5mC coincides with the appearance of 5hmC, suggesting that TET is involved in the rapid disappearance of 5mC^{62–65}. Interestingly, 5fC and 5caC can also be detected in the paternal chromosome, although the significance of this observation is still unknown⁶⁶ (Fig. 3a). Knockdown or targeted deletion of the gene encoding Tet3 — the only highly expressed TET protein in the zygote — abolished the loss of 5mC and the generation of 5hmC, indicating that Tet3 is responsible for the oxidation of 5mC in this context^{62,65}. Immunostaining and sequencing studies have shown that, after the two pronuclei fuse, both the maternal genome containing 5mC and the paternal genome with Tet3-generated 5hmC are diluted in a replication-dependent manner^{63,66,67}. Thus, Tet3 seems to mediate active demethylation of the paternal genome through active oxidation of 5mC followed by passive dilution, resulting in restoration of unmodified C. The reason why the male genome undergoes an additional oxidation step is currently unknown. However, the process is likely to be biologically important, because female mice depleted of Tet3 in the germ line show reduced fecundity and their heterozygous mutant offspring suffer an increased incidence of developmental failure⁶².

What is the mechanism underlying this asymmetric DNA demethylation? Although factors in the paternal genome that attract Tet3 cannot be ruled out, available data suggest that Tet3 may be actively excluded from the maternal genome. A recent study has shown that the dimethylation of histone H3 lysine 9 (H3K9me2) present predominantly on maternal chromatin provides a binding site for the recruitment of PGC7 (also known as Dppa3 and Stella), which in turn excludes Tet3 from binding to the maternal pronucleus⁶⁸. Interestingly, some imprinted loci on the paternal genome that do not undergo demethylation are also not targeted by TET3 (ref. 62). These imprinted sites show similar hallmarks of H3K9me2 and PGC7, suggesting a potential common mechanism for Tet3 exclusion⁶⁸.

TET proteins in PGC reprogramming

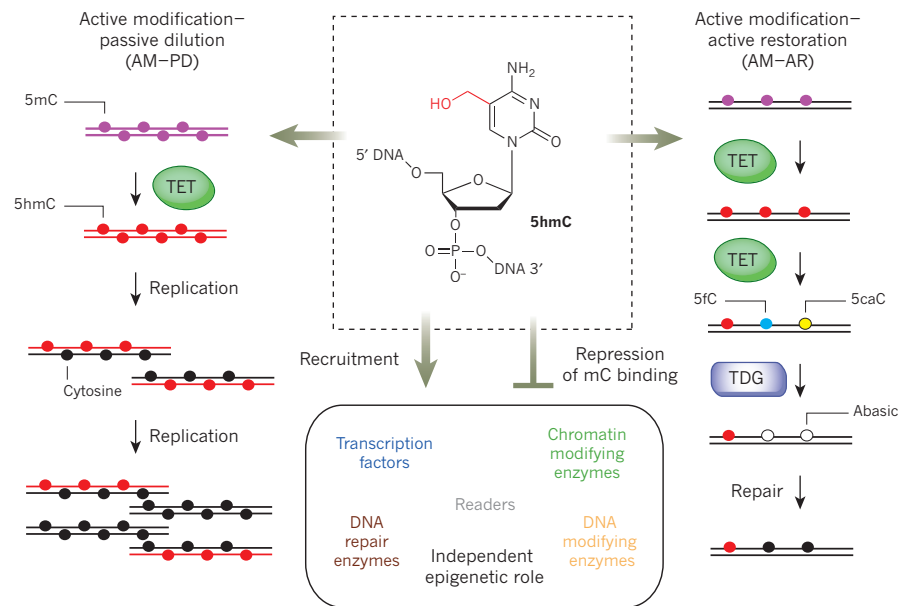
After methylation patterns are established in the embryo, the specialized group of PGCs undergo a further, complex epigenetic reprogramming process that includes erasure of genome-wide DNA methylation patterns¹¹ (Fig. 3b). Although this process is largely believed to be an active process, careful studies using complementary immunostaining and sequencing techniques have revealed that both passive and active processes contribute to the global loss of 5mC^{69–72}. After initial passive dilution of 5mC, 5hmC subsequently accumulates actively and is then lost in an apparent replication-dependent manner^{69,71} (Fig. 3b). Just as in early embryonic development, specific loci can deviate from these global patterns, and these differentially methylated loci can persist even into mature oocytes^{69–71}.

Although both Tet1 and Tet2 are expressed during PGC reprogramming, only Tet1 is upregulated in reprogramming germ cells^{69,73}. However, targeted deletion in mice^{73,74} or knockdown in ES cells followed by *in vitro* PGC differentiation⁷⁵ revealed that Tet1 does not affect global

BOX 1

Roles of 5hmC in DNA demethylation

With the discovery of TET, 5hmC has taken on a new central role in epigenetics. The base sits at an important branch point, with at least three potential outcomes. 5hmC could have an independent epigenetic role related to the base's interaction with chromatin-associated proteins, either through direct recruitment by 5hmC or through disruption of 5mC-specific interactions. 5hmC also can feed into two different pathways that we define as active demethylation (see Perspectives). The pathway involving active modification and passive dilution (AM-PD) (Box Fig., left) is marked by 5mC conversion to 5hmC (or higher oxidized species), replication-dependent depletion of the modified base and reversion to unmodified C. In the active restoration pathway (AM-AR) (Box Fig., right), iterative oxidation by TET can yield 5fC or 5caC, which can be excised by TDG to generate abasic sites as part of the DNA repair pathway that ultimately regenerates unmodified C.



DNA demethylation. Nevertheless, loss of function of Tet1 does impact locus-specific DNA demethylation, particularly at meiotic genes⁷³. In addition, although *Tet2* knockout alone does not lead to any PGC phenotype^{76–78}, demethylation of some imprinted loci is affected in *Tet1* and *Tet2* double knockout mice⁷⁹. Thus, although there is some consensus for a function for TETs in PGCs, further studies are needed to clarify the exact contribution of Tet1 and Tet2 and their possible redundancy in shaping the PGC methylome.

TET proteins in stem cells

ES cells are a model for understanding demethylation dynamics, because maintenance of ES cells is associated with a distinct methylation pattern that supports expression of pluripotency factors while silencing lineage-specification factors. Both Tet1 and Tet2 are expressed in mouse ES cells^{18,80}. TET proteins are probably part of the pluripotency regulatory circuit and may act by directly regulating expression of key ES cell transcription factors^{80,81}. In a series of studies consistent with this hypothesis, short-hairpin-mediated knockdown of *Tet1* alone^{18,82} or in combination with *Tet2* (ref. 80) resulted in a defect in ES cell maintenance, as well as skewed differentiation toward trophectoderm and primitive endoderm. However, some ambiguity remains about the role of TET in ES cell maintenance, given that other knockdown studies do not result in similar phenotypes⁸³ and mice deficient in Tet1 can be derived from *Tet1*-knockout ES cells⁷⁴. The use of different cell lines and culture conditions may contribute to these different results, although potential off-target activity of short hairpin RNAs cannot be ruled out.

Genome-wide and single-base resolution methods have been adapted to discriminate between modified C bases in ES cells. A consensus of these studies (see ref. 84 for a review) has demonstrated a probable regulatory role for 5hmC with particular enrichment at transcribed gene bodies, bivalent and silent promoters, and distal *cis*-regulatory elements. More strikingly, recent genome-wide mapping in ES cells has pointed to the functional relevance of TDG, 5fC and 5caC. Using an immunoprecipitation approach in *Tdg*-deficient ES cells, a significant enrichment of 5fC and 5caC was observed in non-repetitive

regions, particularly at distal regulatory elements⁸⁵. A second study that used a chemical-labelling pull-down approach for detection of 5fC demonstrated 5fC enrichment in enhancer regions⁸⁶. These studies strongly suggest that dynamic C modification involving TDG-mediated 5fC and 5caC removal takes place widely in mouse ES cells.

Although the exact function of the TET proteins in ES cells needs further study, several recent publications are supportive of a role for TET in reprogramming of somatic cells to generate induced pluripotent stem cells (iPSCs) (see Review by Apostolou and Hochedlinger⁸⁷ in this Insight). For example, at the early stage of transduction with the transcription factors Oct4, Klf4, Sox2 and c-Myc (collectively referred to as OKSM), Tet2 is recruited to the *Nanog* and *Esrrb* loci to activate their transcription⁸⁸. In addition, both Tet1 and Tet2 can associate with Nanog and facilitate iPSC generation in an enzymatic activity-dependent manner⁸⁹. Remarkably, Tet1 overexpression can not only enhance reprogramming efficiency by promoting demethylation and reactivation of Oct4, but can also replace Oct4 in the iPSC reprogramming cocktail⁹⁰. Furthermore, beyond reprogramming mediated by OKSM, Tet1 and Tet2 seem to have distinct roles in reprogramming mediated by fusion of somatic cells to pluripotent cells⁹¹.

Locus-specific active demethylation in somatic cells

The key players in the C modifying pathway have also been implicated in locus-specific demethylation, independent of replication. For example, TDG has been observed at loci at which rapid cycling of C and 5mC is associated with hormonal^{12,13} or cytokine-mediated¹⁴ regulation, and TET has been associated with demethylation in the post-mitotic adult brain⁴⁹. These studies imply that active demethylation with TET and TDG may be operational when transcriptional control must be modulated in the absence of DNA replication. However, we still have much more to learn at the level of individual promoters.

DNA demethylation in cancer

Aberrant DNA methylation is a prominent feature of cancer cells⁹², raising the possibility that demethylation pathways may contribute to cancer development⁹³. *TET1* was initially identified owing

to its fusion to *MLL* (also known as *KMT2A*) in patients with acute myeloid leukaemia²⁴, and inactivating *TET2* mutations have since been demonstrated to be frequent lesions in myeloid lineage malignancies^{94,95}. Interestingly, these same myeloid-lineage conditions are susceptible to therapy aimed at inhibiting DNA methylation⁹⁶. Further supporting a role for Tet2 in normal haematopoiesis, mouse models have shown that the enzyme is a crucial regulator of self-renewal and differentiation in haematopoietic stem cells^{76–78,97}. Although most studies have focused on haematological malignancies, downregulation of TET expression has been observed in human breast, liver, lung, pancreatic and prostate cancers⁹⁸. Despite discrepancies in the levels of 5mC in these various settings, TET mutations are consistently associated with a decrease in 5hmC, which has been suggested as a potential diagnostic biomarker^{98,99}. Regarding the other players in demethylation, although the relevance of 5fC and 5caC in cancer has not yet been explored, TDG has also been implicated in various cancers⁴⁵. It remains to be established if this association is due to TDG's role in mismatch repair or in active DNA demethylation.

Interestingly, in acute myeloid leukaemia, TET2 mutations were found to be mutually exclusive with a neomorphic mutation in the isocitrate dehydrogenase genes *IDH1* and *IDH2* (ref. 100). Wild-type *IDH1* and *IDH2* catalyse the conversion of isocitrate to α -KG, the cofactor for the TET and histone demethylase family of oxygenase enzymes³². The neomorphic mutation in *IDH1* and *IDH2* leads to the production of α -hydroxyglutarate, an oncometabolite that can competitively inhibit these α -KG-dependent enzymes¹⁰¹. These studies suggest that neomorphic *IDH1* and *IDH2* mutations may alter DNA methylation patterns by recapitulating TET2 mutations, although alternative mechanisms have also been postulated¹⁰².

Perspective and open questions

Across various physiological developmental niches, non-physiological settings such as iPSCs, and even pathological settings such as cancer, loss of TET proteins and 5hmC is associated with dysregulated DNA methylation. These biological studies, on the heels of a series of transformative biochemical discoveries on TET and TDG, have established 5hmC as a key intermediate in active DNA demethylation.

Revisiting the definition of active demethylation

Recent advances require the classical definitions of passive and active DNA demethylation to be revisited. As we have noted, passive demethylation seems to be best suited for describing the replication-dependent dilution of 5mC only, as this pathway does not involve any active enzymatic processes that alter the base itself. Given our current understanding, active demethylation involving TET is best viewed as two pathways, both of which initially involve active modification (AM) of 5mC to generate 5hmC. This base can be further processed through either passive dilution (PD) to regenerate unmodified C through DNA replication, or active restoration (AR) through further enzymatic modification (Box 1). This framework should also be fitting for other potential pathways for demethylation, such as a 5mC deamination–BER pathway, which would be described as an AM–AR active demethylation pathway.

AM–AR has the advantage of achieving rapid conversion of 5mC to unmodified C, yet the pathway also poses the potential risk of genomic damage given the involvement of DNA breaks in BER. By contrast, the dependence of AM–PD on replication means that functions that might be associated with 5mC modification are quickly achieved, whereas reversion to unmodified C awaits DNA replication. AM–AR therefore seems particularly well suited to locus-specific demethylation processes that require a rapid response to environmental stimuli, whereas AM–PD might be better suited to developmental processes in which cellular replication is tied to lineage specification, such as preimplantation development and PGC reprogramming. With this framework, future studies can evaluate

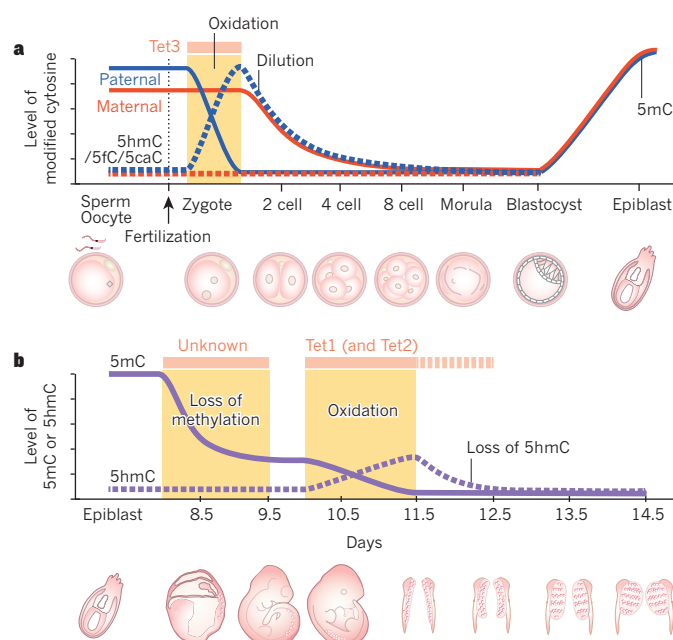


Figure 3 | DNA methylation dynamics in pre-implantation embryos and primordial germ cells.

a, Dynamics of 5mC and its oxidation products in pre-implantation embryos. Although the maternal DNA goes through passive demethylation, the paternal genome is demethylated in two steps. Tet3 first oxidizes the 5mC in the paternal genome, and the oxidation products are then diluted through a replication-dependent process. For clarity, although the absolute levels of 5hmC, 5fC and 5caC differ, the bases are schematically shown together (dotted line) given that their increase and subsequent depletion follow similar patterns. DNA methylation patterns are re-established by *de novo* DNMTs at the blastocyst stage. **b**, Illustration of the 5mC and 5hmC dynamics in primordial germ cells (PGCs) during their reprogramming. DNA demethylation in PGCs goes through three stages: loss of bulk DNA methylation in a Tet-independent manner; oxidation of remaining 5mC to 5hmC by Tet1 and potentially Tet2 proteins; and loss of 5hmC through replication-dependent passive dilution. 5fC and 5caC are not shown in this panel because no dynamic change in their levels was observed by immunostaining^{69–72}. Figure scale is shown in embryonic days post-fertilization.

the theoretical risks and benefits of the AM–PD and AM–AR pathways and better delineate the cellular context in which either or both pathways are active.

Regulation of the demethylation pathway

Viewing C modification as a series of step-wise modifications (Fig. 2) prompts the key question: what regulates stalling at various intermediates in the pathway or progression through the cycle?

5hmC is significantly more prevalent than 5fC and 5caC⁸⁴. Given that TET enzymes can iteratively oxidize, it remains unclear what factors dictate that the modification pathway halts at 5hmC. Stalling of the pathway at 5hmC could be regulated through modulating TET's accessibility to 5hmC, through either post-translational modifications or interaction with protein partners. Alternatively, stalling at 5hmC could be regulated at a biochemical level through altered enzyme kinetics. In this regard, a crucial question that remains unresolved refers to TET's relative ability to oxidize 5mC, 5hmC and 5fC. The basal reactivity of TET with each of these substrates and regulation of its substrate preferences will need to be addressed. Structural insights into the TET catalytic domain could prove key to deciphering regulatory mechanisms that govern iterative oxidation.

At the next stage of the pathway, do 5fC and 5caC have distinctive roles and what, if any, significance do they have beyond serving as intermediates in demethylation? 5fC and 5caC are similar marks in that they both result from iterative oxidation and both can be excised by TDG, yet they

could have different roles. Indeed, TDG shows a higher affinity for 5fC⁶⁰ and different mechanisms may be used in excision of 5fC and 5caC³¹. Given the relative scarcity of 5fC and 5caC, it is not clear if these bases are simply intermediates in an active demethylation pathway, or if they have functionally significant interactions with genomic 'readers'^{103,104}. Further efforts to perturb TDG function and to localize 5fC and 5caC in the genome in high resolution and at specific genomic loci should help to resolve some of these issues.

When the pathway is viewed as a complete cycle of C methylation, oxidation and repair, it immediately begs the question: does recurrent cycling of this pathway occur? Rapid cycling between C and 5mC has been observed at some promoters^{12,13} and disruption of TDG leads to accumulation of intermediates in the pathway^{85,86}. However, bona fide evidence of multiple cycling events through an AM-AR pathway at a single locus has yet to be shown.

Finally, although we have mainly emphasized the emerging role of TET and TDG, numerous other DNA modifying enzymes that can recognize modified C bases will undoubtedly influence the physiological function of the dynamic demethylation pathway. An active search for potential 5caC decarboxylase enzymes, further studies to elucidate the relevance of AID/APOBEC deamination of 5mC, the role of DNMTs in a 'reverse' reaction, and interactions of modified C bases with DNA-binding proteins are just a few areas in which advances may come in the next phase of discovery.

Oxidative modifications of 5mC and related repair mechanisms have expanded the possibilities by which the genome can retain great flexibility while maintaining the integrity of its coding information. A few years ago, it would have been hard to imagine how much our knowledge of active DNA demethylation would change, and we anticipate that the years ahead will be marked by many more exciting discoveries regarding the role of dynamic regulation of DNA methylation in development, gene regulation and genome stability. ■

Received 2 May; accepted 6 August 2013.

- Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74**, 481–514 (2005).
- Klose, R. J. & Bird, A. P. Genomic DNA methylation: the mark and its mediators. *Trends Biochem. Sci.* **31**, 89–97 (2006).
- Bestor, T. H. & Bourc'his, D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb. Symp. Quant. Biol.* **69**, 381–387 (2004).
- Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.* **33** (suppl.), 245–254 (2003).
- Wu, S. C. & Zhang, Y. Active DNA demethylation: many roads lead to Rome. *Nature Rev. Mol. Cell Biol.* **11**, 607–620 (2010).
- Ooi, S. K. & Bestor, T. H. The colorful history of active DNA demethylation. *Cell* **133**, 1145–1148 (2008).
- Surani, M. A. & Hajkova, P. Epigenetic reprogramming of mouse germ cells toward totipotency. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 211–218 (2010).
- Jenkins, T. G. & Carrell, D. T. Dynamic alterations in the paternal epigenetic landscape following fertilization. *Front. Genet.* **3**, 143 (2012).
- Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
- Oswald, J. et al. Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
- Hackett, J. A., Zyllicz, J. J. & Surani, M. A. Parallel mechanisms of epigenetic reprogramming in the germline. *Trends Genet.* **28**, 164–174 (2012).
- Kangaspeska, S. et al. Transient cyclical methylation of promoter DNA. *Nature* **452**, 112–115 (2008).
- Métivier, R. et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature* **452**, 45–50 (2008).
- Thilainadesan, G. et al. TGF- β -dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol. Cell* **46**, 636–649 (2012).
- Bhutani, N., Burns, D. M. & Blau, H. M. DNA demethylation dynamics. *Cell* **146**, 866–872 (2011).
- Nabel, C. S., Manning, S. A. & Kohli, R. M. The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. *ACS Chem. Biol.* **7**, 20–30 (2012).
- Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009). **This manuscript reports on the discovery of TET enzymes and their biochemical ability to convert 5mC to 5hmC.**
- Ito, S. et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010). **This report describes the discovery of a role for TET enzymes in 5mC oxidation within ES cells.**
- Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).
- Borst, P. & Sabatini, R. Base J: discovery, biosynthesis, and possible functions. *Annu. Rev. Microbiol.* **62**, 235–251 (2008).
- Liu, C. K., Hsu, C. A. & Abbott, M. T. Catalysis of three sequential dioxygenase reactions by thymine 7-hydroxylase. *Arch. Biochem. Biophys.* **159**, 180–187 (1973).
- Smiley, J. A., Kundracik, M., Landfried, D. A., Barnes, V. R. S. & Axheim, A. A. Genes of the thymidine salvage pathway: thymine-7-hydroxylase from a *Rhodotorula glutinis* cDNA library and iso-orotate decarboxylase from *Neurospora crassa*. *Biochim. Biophys. Acta* **1723**, 256–264 (2005).
- Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710 (2009).
- Ono, R. et al. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* **62**, 4075–4080 (2002).
- Penn, N. W., Suwalski, R., O'Riley, C., Bojanowski, K. & Yura, R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem. J.* **126**, 781–790 (1972).
- Globisch, D. et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS ONE* **5**, e15367 (2010).
- Song, C. X. et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature Biotechnol.* **29**, 68–72 (2011).
- Ito, S. et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011). **This report describes the discovery that TET can iteratively oxidize 5mC to 5hmC, 5fC and 5caC.**
- He, Y. F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011). **This manuscript shows the generation of 5caC by TET and its excision by TDG.**
- Williams, R. T. & Wang, Y. A density functional theory study on the kinetics and thermodynamics of N-glycosidic bond cleavage in 5-substituted 2'-deoxycytidines. *Biochemistry* **51**, 6458–6462 (2012).
- Hashimoto, H., Zhang, X. & Cheng, X. Selective excision of 5-carboxylcytosine by a thymine DNA glycosylase mutant. *J. Mol. Biol.* **425**, 971–976 (2013).
- Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem. Sci.* **36**, 7–18 (2011).
- Xu, Y. et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* **151**, 1200–1213 (2012).
- Zhang, H. et al. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.* **20**, 1390–1393 (2010).
- Ko, M. et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **497**, 122–126 (2013).
- Schiesser, S. et al. Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew. Chem. Int. Edn Engl.* **51**, 6516–6520 (2012).
- Lutkeviciute, Z., Lukinavicius, G., Masevicius, V., Daujotyte, D. & Klimasauskas, S. Cytosine-5-methyltransferases add aldehydes to DNA. *Nature Chem. Biol.* **5**, 400–402 (2009).
- Niehrs, C. & Schafer, A. Active DNA demethylation by Gadd45 and DNA repair. *Trends Cell Biol.* **22**, 220–227 (2012).
- Ma, D. K., Guo, J. U., Ming, G. L. & Song, H. DNA excision repair proteins and Gadd45 as molecular players for active DNA demethylation. *Cell Cycle* **8**, 1526–1531 (2009).
- Fromme, J. C. & Verdine, G. L. Base excision repair. *Adv. Protein Chem.* **69**, 1–41 (2004).
- Hajkova, P. et al. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78–82 (2010).
- Gehring, M., Reik, W. & Henikoff, S. DNA demethylation by DNA repair. *Trends Genet.* **25**, 82–90 (2009).
- Wossidlo, M. et al. Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877–1888 (2010).
- Zhu, J. K. Active DNA demethylation mediated by DNA glycosylases. *Annu. Rev. Genet.* **43**, 143–166 (2009).
- Dalton, S. R. & Bellacosa, A. DNA demethylation by TDG. *Epigenomics* **4**, 459–467 (2012).
- Cortázar, D. et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**, 419–423 (2011).
- Cortellino, S. et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67–79 (2011).
- Cortázar, D., Kunz, C., Saito, Y., Steinacher, R. & Schar, P. The enigmatic thymine DNA glycosylase. *DNA Repair (Amst.)* **6**, 489–504 (2007). **References 47 and 48 demonstrate that deletion of TDG is associated with embryonic lethality and epigenetic abnormalities.**
- Guo, J. U., Su, Y., Zhong, C., Ming, G. L. & Song, H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423–434 (2011).
- Bhutani, N. et al. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042–1047 (2010).
- Popp, C. et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
- Kumar, R. et al. AID stabilizes stem-cell phenotype by removing epigenetic

- memory of pluripotency genes. *Nature* **500**, 89–92 (2013).
53. Rai, K. *et al.* DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**, 1201–1212 (2008).
 54. Barreto, G. *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**, 671–675 (2007).
 55. Jin, S. G., Guo, C. & Pfeifer, G. P. GADD45A does not promote DNA demethylation. *PLoS Genet.* **4**, e1000013 (2008).
 56. Wong, E. *et al.* Mbd4 inactivation increases C-T transition mutations and promotes gastrointestinal tumor formation. *Proc. Natl Acad. Sci. USA* **99**, 14937–14942 (2002).
 57. Nabel, C. S. *et al.* AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nature Chem. Biol.* **8**, 751–758 (2012).
 58. Rangam, G., Schmitz, K. M., Cobb, A. J. & Petersen-Mahrt, S. K. AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS ONE* **7**, e43279 (2012).
 59. Bennett, M. T. *et al.* Specificity of human thymine DNA glycosylase depends on N-glycosidic bond stability. *J. Am. Chem. Soc.* **128**, 12510–12519 (2006).
 60. Maiti, A. & Drohat, A. C. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* **286**, 35334–35338 (2011).
 61. Zhang, L. *et al.* Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nature Chem. Biol.* **8**, 328–330 (2012).
 62. Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
- This report demonstrates that TET3 is responsible for paternal genome 5mC oxidation in zygotes.**
63. Inoue, A. & Zhang, Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194 (2011).
- This report demonstrates that 5hmC generated in zygotes is passively diluted during preimplantation development.**
64. Iqbal, K., Jin, S. G., Pfeifer, G. P. & Szabo, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* **108**, 3642–3647 (2011).
 65. Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Commun.* **2**, 241 (2011).
 66. Inoue, A., Shen, L., Dai, Q., He, C. & Zhang, Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.* **21**, 1670–1676 (2011).
 67. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
 68. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
 69. Hackett, J. A. *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).
 70. Seisenberger, S. *et al.* The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).
 71. Yamaguchi, S. *et al.* Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res.* **23**, 329–339 (2013).
 72. Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. & Saitou, M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**, 340–353 (2013).
 73. Yamaguchi, S. *et al.* Tet1 controls meiosis by regulating meiotic gene expression. *Nature* **492**, 443–447 (2012).
 74. Dawlaty, M. M. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175 (2011).
 75. Vincent, J. J. *et al.* Stage-specific roles for Tet1 and Tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell* **12**, 470–478 (2013).
 76. Li, Z. *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**, 4509–4518 (2011).
 77. Moran-Crusio, K. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
 78. Quivoron, C. *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* **20**, 25–38 (2011).
 79. Dawlaty, M. M. *et al.* Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–323 (2013).
 80. Koh, K. P. *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**, 200–213 (2011).
 81. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
 82. Freudenberger, J. M. *et al.* Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Res.* **40**, 3364–3377 (2012).
 83. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
 84. Song, C. X., Yi, C. & He, C. Mapping recently identified nucleotide variants in the genome and transcriptome. *Nature Biotechnol.* **30**, 1107–1116 (2012).
- This review details the novel methods that have been developed to quantify and sequence modified C bases in the genome.**
85. Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013).
 86. Song, C. X. *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* **153**, 678–691 (2013).
 87. Apostolou, E. & Hochedlinger, K. Chromatin dynamics during somatic cell reprogramming. *Nature* **502**, 462–471 (2013).
 88. Doege, C. A. *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**, 652–655 (2012).
 89. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).
 90. Gao, Y. *et al.* Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
 91. Piccolo, F. M. *et al.* Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell* **49**, 1023–1033 (2013).
 92. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nature Rev. Cancer* **11**, 726–734 (2011).
 93. Cimmino, L., Abdel-Wahab, O., Levine, R. L. & Aifantis, I. TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* **9**, 193–204 (2011).
 94. Delhommeau, F. *et al.* Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
 95. Langemeijer, S. M. *et al.* Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nature Genet.* **41**, 838–842 (2009).
 96. Quintás-Cardama, A., Santos, F. P. & Garcia-Manero, G. Therapy with azanucleosides for myelodysplastic syndromes. *Nature Rev. Clin. Oncol.* **7**, 433–444 (2010).
 97. Ko, M. *et al.* Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl Acad. Sci. USA* **108**, 14566–14571 (2011).
 98. Yang, H. *et al.* Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene* **32**, 663–669 (2013).
 99. Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
 100. Figueroa, M. E. *et al.* Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553–567 (2010).
 101. Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).
 102. Losman, J. A. *et al.* (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science* **339**, 1621–1625 (2013).
 103. Spruijt, C. G. *et al.* Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013).
 104. Mellén, M., Ayata, P., Dewell, S., Kriacounis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).

Acknowledgements We would like to thank S. Yamaguchi for preparing Figure 3; D. Crawford, A. Inoue, C. Nabel, E. Schutsky and S. Yamaguchi for their helpful comments. We apologize to the people whose work cannot be cited due to space limitations. Our DNA methylation-related work is supported by the NIH (U01DK089565 and GM068804 to Y.Z., and K08-AI089242 to R.M.K.), the Rita Allen Foundation (R.M.K.) and HHMI (Y.Z.). Y.Z. is an investigator of the HHMI.

Author Information Reprints and permission information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/ezvkfv. Correspondence should be addressed to R.K. (rkohli@upenn.edu) or Y. Z. (yzhang@genetics.med.harvard.edu).

Chromatin proteins and modifications as drug targets

Kristian Helin^{1,2,3} & Dashyant Dhanak⁴

A plethora of groundbreaking studies have demonstrated the importance of chromatin-associated proteins and post-translational modifications of histones, proteins and DNA (so-called epigenetic modifications) for transcriptional control and normal development. Disruption of epigenetic control is a frequent event in disease, and the first epigenetic-based therapies for cancer treatment have been approved. A generation of new classes of potent and specific inhibitors for several chromatin-associated proteins have shown promise in preclinical trials. Although the biology of epigenetic regulation is complex, new inhibitors such as these will hopefully be of clinical use in the coming years.

Epigenetics is defined as heritable traits that are not linked to changes in the DNA sequence; however, in broader terms, epigenetics is used to describe the mechanisms by which chromatin-associated proteins and post-translational modifications (PTMs) of histones regulate transcription. Although all cells within an organism contain the same DNA, epigenetic regulators and transcription factors organize the genome into accessible and closed regions, which ensure the correct transcriptional program in a given cell type. Thus, epigenetic regulation is important for maintaining cell identity and is implicated in fundamental processes such as proliferation, development, differentiation and genome integrity. Epigenetic gene regulation can be mediated through DNA methylation, nucleosome remodelling, exchange of histone variants and PTMs of the histones (Box 1). Histones can be modified at specific amino acids with a diverse set of chemical modifications, such as phosphorylation, acetylation, methylation, ubiquitination or SUMOylation^{1,2}. Research in the past decade has led to a better understanding of the significance of these PTMs. In particular, this progress has been achieved through the identification of chromatin-associated proteins that catalyse, recognize and remove the specific modification (Box 1), the generation of high affinity antibodies specific for the PTM, genome-wide location analysis and genetic studies.

Deregulation of epigenetic control is a common feature of a number of diseases, including brain disorders and cancer³. The involvement of DNA methylation in cancer has been appreciated for a number of years, and the approval of the first drugs targeting DNA methylation is a hallmark for epigenetic-based therapies. The two approved drugs, azacitidine (5-azacytidine) and decitabine (5-aza-2'-deoxycytidine), are nucleoside analogues and irreversible inhibitors of the DNA methyltransferase enzymes DNMT1 and DNMT3. They are currently used as first-line treatments for patients with myelodysplastic syndrome^{4,5}. Shortly after the approval of the two DNA methylation inhibitors, the two histone deacetylase (HDAC) inhibitors suberoylanilide hydroxamic acid (SAHA) and romidepsin (depsipeptide or FK228) were approved for the treatment of refractory cutaneous T-cell lymphoma^{6,7}. Although the introduction of these drugs in the clinic has been a tremendous success for the field, a number of scientific challenges remain. Despite many years of research, we do not understand exactly how and why these drugs work. For HDAC inhibitors, acetylation is in general increased following drug treatment; however, data demonstrating a correlation between HDAC activity and therapeutic index is still lacking. Similarly,

so far there is no established gene expression signature or profile that can predict whether a patient will benefit from the use of HDAC inhibitors. The picture is very similar for DNMT inhibitors. Although these molecularly targeted drugs have the potential to revert the epigenetic modification and have been shown to lead to global hypomethylation, we do not know their precise mechanism of action. For both classes of drug, the lack of reliable molecular biomarkers for predicting either clinical activity or resistance is a serious drawback, limiting clinicians' ability to achieve the vision of 'personalized medicine'. Despite a large number of clinical trials, the use of the four drugs is so far limited to specific haematological cancers.

Recently, the use of next-generation sequencing technologies on DNA isolated from primary tumours has revealed a high frequency of somatic mutations in genes coding for chromatin-associated proteins that are known to regulate DNA methylation patterns, histone PTMs and chromatin remodelling (see ref. 8 for a review). Strikingly, the discovery that patients with leukaemia often have mutations in genes such as *TET2*, *IDH1*, *IDH2* and *DNMT3A*, which are all involved in regulating DNA methylation patterns, might provide insight into why patients with leukaemia show a significant response to DNA methylation inhibitors, and could hold promise for future patient stratification strategies. In fact, the lack of genetic data to support the role of chromatin-associated proteins in cancer has been a major obstacle for the development of patient-specific targeted therapies. This has drastically changed with the recent findings that chromatin-associated proteins often show aberrant expression in cancer as a result of translocations or genetic amplifications, and by the discovery that they carry specific somatic mutations.

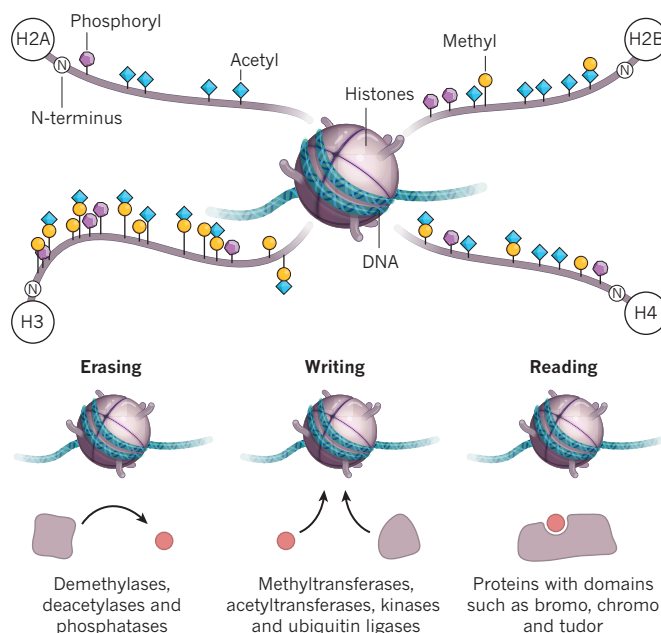
In this Review, we will focus on the recent advances made by the scientific and pharmaceutical communities to develop highly potent and specific inhibitors to chromatin-associated proteins (Table 1). These represent several new classes of therapeutic targets and, as we will exemplify, recent results have shown the feasibility of developing specific inhibitors to histone methyltransferases (HMTs), histone demethylases and domains required for the binding of protein complexes to specific histone modifications. This is a very exciting time for the field, in which the combination of knowledge regarding the role of chromatin-associated proteins in disease and the development of potential new classes of epigenetic drugs will hopefully lead to molecularly targeted and lower toxicity therapies with a clear genetic marker or markers for patient stratification.

¹Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark. ²Centre for Epigenetics, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark. ³The Danish Stem Cell Center, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark. ⁴Janssen Research & Development, 1400 McKean Road, PO Box 776, Spring House, Pennsylvania 19477, USA.

BOX 1

The role of DNA and histone PTMs

DNA is wrapped around histones (H2A, H2B, H3 and H4) to form nucleosomes. Nucleosomes are further compacted to form condensed chromatin. The compaction of DNA is in part regulated through post-translational modifications (PTMs) of the histone tails, which protrude from nucleosomes. Epigenetic regulators can in popular terms be divided into erasers, writers or readers of PTMs. The erasers, such as histone deacetylases and histone demethylases, remove the PTMs and prepare the histones for other modifications. The writers comprise enzymes such as histone acetylases, kinases, DNA and histone methyltransferases and ubiquitin ligases. The writers catalyse the PTMs on the DNA or the proteins, and may impose epigenetic heritability such as DNA methylation through copying and maintaining the modification. Other modifications, such as histone acetylation, respond rapidly to environmental stimuli and are therefore more dynamic. Readers of the post-translational modification include proteins with specific domains, such as bromo-, chromo-, tudor-, MBT-, PWWP-, WD40- and PHD-domains, which bind to the specific modification. The readers, which are often found in large protein complexes, interpret the modification and impose changes in chromatin structure.



Targeting histone methyltransferases

An association between histone hypermethylation, transcriptional regulation and the cancer phenotype has spurred efforts to develop specific, small molecule inhibitors of the methyltransferase enzymes involved in histone lysine and arginine methylation. The family of HMTs (or more accurately, protein methyltransferases; that is, protein arginine methyltransferases (PRMT) and protein lysine methyltransferases (KMTs)) encompasses over 60 different proteins that sequentially transfer a methyl group from the cofactor *S*-adenosylmethionine (SAM) to the terminal amine of specific substrate lysine and/or arginine residues. With the notable exception of the HMT DOT1L (see later), the catalytic transfer of a methyl group from SAM occurs within a conserved SET domain, which accommodates the cofactor and peptide substrates in a conformation conducive for an SN2 transfer reaction generating *S*-adenosylhomocysteine (SAH) and the methylated histone side chain as products (Fig. 1). Detailed structural determinations of multiple SET-domain-containing HMTs have been carried out to support this mechanistic rationale for the methyl transfer event with a detailed analysis of binding modes of cofactor and/or peptide substrates to allow the rational design of selective inhibitors. An understanding of exactly how the degree of histone lysine methylation modulates transcription remains to be attained, but the need for the coordinated recruitment of methylation-sensitive proteins to transcriptional complexes offers one plausible hypothesis. Interestingly, the HMTs have also been reported^{9–11} to act on various non-histone protein substrates to regulate their functions. However, the relative contributions of the histone compared with non-histone action of HMTs are not well understood and continue to be an area of active investigation.

In the context of cancer, the discovery of genetic alterations in HMTs in several different tumour types^{12–14} has undoubtedly attracted much attention and provided additional support for the importance of epigenetic deregulation in a disease that is widely considered to be genetically driven. In some cases (such as the methyltransferase EZH2, discussed later), heterozygous point mutations in the catalytic SET domain lead to a gain of function of the wild-type enzyme^{15,16}, favouring trimethylation and the silencing of tumour suppressor genes and/or differentiation-specific genes. Similarly, in other cancers (such as, increased expression of NSD2 in multiple myeloma) chromosomal translocations result in increased

expression of the methyltransferases, again leading to aberrant transcription and proliferation¹⁷. Conversely, lysine methylation induced by the HMT DOT1L results in sustained expression of several genes required for leukaemogenesis. Therefore, small molecule inhibitors, of for instance EZH2 or DOT1L, should be able to reduce or eliminate the site-specific lysine methylation introduced by the HMTs and reverse the oncogenic state (see later).

DOT1L

Chromosomal translocations are relatively common in various haematopoietic malignancies and can be associated with aggressive or poorly responsive disease. In leukaemia that involves rearrangement of the *MLL* (also known as *KMT2A*) gene, translocation leads to fusions with more than 50 different protein partners including ENL, ELL, AF4 and AF9 (ref. 18) (Fig. 2a). The resulting fusion complexes recruit DOT1L, which specifically methylates the core histone H3 residue lysine 79 (H3K79) and contributes to transcriptional activation of *HOXA10*, *MEIS1* and other genes required for leukaemia initiation¹⁹. DOT1L lacks the SET domain that is commonly present in other lysine methyltransferases but nonetheless can readily catalyse the transfer of one, two or three methyl groups to the ϵ -NH₂ group of H3K79. In a crucial paper from the Armstrong laboratory²⁰, deletion of DOT1L in *MLL*-rearranged cell lines and subsequently in *in vivo* mouse studies directly demonstrated the role of the enzyme not only in introducing the H3K79 methyl mark, leading to a concomitant increase in gene expression, but also in the development of the leukaemia.

Given the significant role of DOT1L in *MLL*-rearranged leukaemia, inhibitors of its H3K79 methyltransferase activity have been aggressively pursued as potential therapeutics. EPZ004777, a SAM-competitive pyrrolopyrimidine derivative (Fig. 2b) was designed²¹ to mimic both SAM and the reaction product SAH while also taking advantage of potential hydrophobic interactions available in the binding vicinity. The compound is an extremely potent and remarkably selective SAM-competitive inhibitor of the enzyme. In *MLL*-rearranged cell lines, EPZ004777 reduces global H3K79me2 levels, blocks the expression of *MLL*-fusion target genes and has antiproliferative activity²¹. Consistent with a targeted mechanism of action, only cell lines with an *MLL* gene fusion were

sensitive to the DOT1L inhibitor whereas non-rearranged lines remained unaffected. Regardless of the measured parameter, the kinetics of cellular response to DOT1L inhibition (and other epigenetic drugs reported so far) is strikingly distinct to the more rapid response usually seen within a few hours with signal transduction modulators (kinase inhibitors) or non-specific chemotherapeutic drugs. Thus, the maximal effect on depletion of the methyl mark is typically seen only after 4–5 days of exposure to the drug. Similarly, significant transcriptional changes occur after 6–8 days and more than 10 days are required to observe an antiproliferative phenotype. Defining and understanding these distinctive characteristics have important implications for the development of these agents because established measures of biomarker-based pharmacodynamic and/or early clinical response may be inappropriate. In addition, prolonged exposure to the drug may be required for efficacy, further highlighting the need for a selective compound with presumably lower propensity for undesirable off-target effects. Encouragingly, in preclinical experiments, EPZ004777 seemed to be well tolerated when given to mice at efficacious doses²¹.

Unfortunately, notwithstanding these attractive attributes, poor pharmacokinetics — including a short plasma half-life — requires EPZ004777 to be administered as a 7 day continuous infusion using surgically implanted mini-osmotic pumps. In a preclinical setting, such studies are readily conducted but can pose significant challenges in clinical studies involving patients with cancer. In an attempt to address these shortcomings, further modifications of the pyrrolopyrimidine core of EPZ004777 have been investigated²² as an approach to designing second-generation DOT1L-targeting drugs. For example, the Structural Genomics Consortium (SGC) has described bromo-deaza-SAH (Fig. 2b) as a convenient DOT1L inhibitor, allowing for the generation of X-ray co-crystal structures and hence the rational design of new analogues with improved properties²³. The recent initiation, by the biotech company Epizyme, of clinical trials to determine the safety and efficacy of the DOT1L inhibitor EPZ-5676 (ref. 24) in patients with MLL leukaemia is highly significant and represents the first human study of a ‘targeted’ HMT inhibitor.

EZH2

The enzyme EZH2 is the catalytic component of the Polycomb protein complex PRC2 and acts as an HMT at H3K27. Importantly, in cell-free systems the EZH2 subunit is only catalytically competent when in a complex with at least two non-enzymatic partners (EED and SUZ12) and moreover in a physiologically relevant, intracellular context, the

complex is known to contain two additional proteins (AEBP2 in complex with either RBBP4 or RBBP7)²⁵ (Fig. 2c).

PRC2 maintains the transcriptional repression of a large number of genes with key regulatory roles in development and differentiation, and PRC2 proteins are required for normal embryonic development²⁵. Pioneering studies from the Chinnaiyan lab have shown an association between increased levels of both EZH2 and H3K27me3 and poor outcomes in metastatic prostate cancer²⁶. In addition, inactivating mutations in UTX, an H3K27 demethylase^{27,28}, are also similarly correlated, suggesting a key role for H3K27 hypermethylation in prostate cancer. Other studies have revealed a similar relationship between elevated levels of EZH2 with silencing of EZH2 target genes and poor prognosis in solid tumours, including breast, kidney and lung^{29–32}. More recently, somatic activating mutations in the SET domain of EZH2 have been identified in follicular lymphoma, and diffuse large B-cell lymphoma (DLBCL), leading to increased H3K27me3 (refs 33–35). Taken together, these findings suggest that misregulation of H3K27me3 levels, through EZH2 overexpression or point mutations, silences target genes that are important to tumour growth and survival and make a compelling case for targeting the enzyme therapeutically. Paradoxically, however, inactivating mutations in EZH2 have also been reported in myelodysplastic syndrome³⁶, raising the potential of a tumour suppressor function for the protein. The role of EZH2 and H3K27 methylation in promoting or inhibiting tumorigenesis and/or maintenance seems therefore to be context dependent and, based on the potential for deleterious effects, suggests caution should be taken in developing chronically administered therapeutic inhibitors. Despite these potential drawbacks, multiple pharmaceutical and biotech company research groups have developed highly potent, selective, small molecule inhibitors of EZH2 (refs 37–39), and other investigators have pursued equally interesting natural-product-based inhibitors⁴⁰.

The medicinal chemistry design of HMT inhibitors has sought to take advantage of the intrinsic affinity of EZH2 for both SAH ($K_i = 7.5 \mu\text{M}$) and lysine-containing substrate mimetics. Hybrid molecules (such as that shown in Fig. 2d) that contain discrete elements of both recognition motifs are modest inhibitors and presumably act as classical bi-substrate inhibitors⁴¹. However, the relatively low permeability of these highly charged compounds might limit their use in cell-based assays and *in vivo*. By contrast, despite being devoid of direct EZH2 inhibitory activity, the structurally related and widely used 3-deazaneplanocin (DZNep; Fig. 2d) has been shown to reactivate indirectly PRC2-silenced genes in cancer cell lines by depleting PRC2 subunits⁴². Unfortunately, this activity does not allow for differentiation of selective catalytic inhibition of EZH2 from more global effects of depleting PRC2, including loss of scaffolding functions, microRNA binding sites and so on, and complicates the interpretation of cellular phenotypes resulting from true inhibition of H3K27 methylation⁴³. Ultimately, the use of DZNep in studies related to investigating the role of EZH2 inhibition in bioassays should be avoided.

High throughput screening of distinct compound libraries by various groups led to the discovery of non-SAM-derived catalytic inhibitors of EZH2. Remarkably, all the screens identified compounds with a pyridone amide motif, indicating a crucial molecular recognition role for functionality. Although these molecules do not resemble SAM, biochemically they are competitive inhibitors of cofactor binding and various three dimensional homology models have been proposed to rationalize how they may mimic the interactions of the natural substrate. Ultimately, detailed structural studies will be needed to determine unequivocally if both occupy the same binding site in EZH2. Despite these uncertainties, extensive chemical modification of the hits identified in high throughput screening to improve affinity and pharmaceutical properties led to the discovery of analogues^{37–39} (such as those shown in Fig. 2d), all of which were highly potent, selective and bioavailable inhibitors of EZH2 in biochemical and cellular assays with *in vivo* antitumour activity in germinal-cell DLBCL with activating EZH2 mutations. Remarkably, these compounds show exquisite selectivity for EZH2 inhibition

Table 1 | Small molecule inhibitors to chromatin-associated proteins

Chromatin-binding protein	Compound
Histone methyltransferases	
DOT1L	EPZ004777 (ref. 21), EPZ-5676 (ref. 24), SGC0946 (ref. 86)
EZH2	GSK126 (ref. 37), GSK343 (refs 87, 88), EPZ005687 (ref. 38), EPZ-6438 (ref. 44), EI1 (ref. 39), UNC1999 (ref. 89)
G9A	BIX01294 (ref. 90), UNC0321 (ref. 91), UNC0638 (ref. 92), NC0642 (ref. 88), BRD4770 (ref. 93)
PRMT3	14u (ref. 94)
PRMT4 (CARM1)	17b (Bristol-Myers Squibb) (refs 95, 96), MethylGene (ref. 97)
Histone demethylases	
LSD1	Tranylcypromine (ref. 62), ORY-1001 (ref. 63)
Bromodomains	
BET	JQ1 (ref. 73), IBET762 (ref. 72), IBET151 (refs 76, 98), PFI-1 (ref. 99)
BAZ2B	GSK2801 (ref. 88)
Chromodomains	
L3MBTL1	UNC669 (ref. 100)
L3MBTL3	UNC1215 (ref. 101)

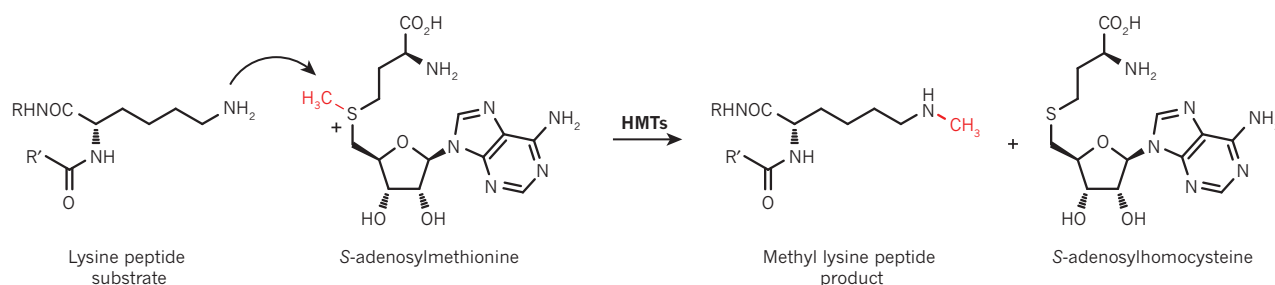


Figure 1 | Mechanism of lysine methylation catalysed by histone lysine methyltransferases. The lysine amino group of the substrate histone polypeptide engages in an SN2 reaction with the activated co-factor S-adenosylmethionine (SAM), resulting in the formation of an N-methylated lysine and S-adenosylhomocysteine (SAH).

(more than 10,000-fold) over most other methyltransferases and can distinguish from EZH1 inhibition (around 100-fold). One of these compounds (EPZ-6438, also known as E7438) has entered human clinical trials and several others are likely to follow shortly, allowing for an assessment of the therapeutic potential of targeting EZH2 in not only lymphoma but also solid tumours with increased levels of H3K27me3. In this context, the recent report of the activity of an EZH2 inhibitor in a preclinical model of paediatric malignant rhabdoid cancer is notable. A subset of these tumours with inactivated SMARCB1 are thought to be dependent on the catalytic activity of EZH2, and in xenograft models were shown to be sensitive to treatment with the EZH2 inhibitor, EPZ-6438 (ref. 44). Interestingly, and as mentioned above, other solid tumours (for example, prostate and breast) have also been associated with drastic upregulation of EZH2 expression but surprisingly, no convincing data has emerged showing activity of catalytic EZH2 inhibitors in these cancers. As with many other new potential therapeutics, the safety profile of EZH2 inhibitors remains to be fully defined but initial observations in prolonged animal studies suggest that they are well tolerated with little or no overt toxicity and EPZ-6438 has been advanced to a phase 1/2 clinical trial in patients with advanced solid tumours or with B-cell lymphomas.

Targeting histone demethylases

Previously, methylation was considered to constitute a permanent and irreversible histone modification that defined epigenetic programs in concert with DNA methylation. However, the discovery of lysine-specific demethylase 1 (LSD1, also known as KDM1A, AOF2, BHC110 and KIAA0601) and later the JmjC-domain-containing lysine demethylase family has completely changed this view (for reviews, see refs 45, 46). LSD1 and its close relative LSD2 (also known as KDM1B and AOF1) belong to the superfamily of flavin adenine dinucleotide (FAD)-dependent monooxidases (Fig. 3a). The two proteins can catalyse the demethylation of H3K4me2 and H3K4me1, and LSD1 has in addition been shown to catalyse the demethylation of H3K9me2 and H3K9me1 as well as a number of non-histone proteins such as p53, DNMT1 and E2F1.

The JmjC-domain family

In contrast to the LSD demethylases, the JmjC-domain-containing demethylases can also demethylate trimethylated lysines. This catalysis involves an oxidative mechanism requiring iron and 2-oxoglutarate as co-factors and probably occurs through direct hydroxylation of the affected methyl group (Fig. 3b)^{45,46}. There are 30 of these JmjC-domain-containing proteins in humans, of which 17 have been shown to be active histone lysine demethylases. Several results have associated the histone lysine demethylases with disease, in particular cancer and brain disorders. For instance, members of the JMJD2 (also known as KDM4) family, which can demethylate H3K9me3 and H3K9me2, and H3K36me3 and H3K36me2 have been found to be overexpressed in squamous cell carcinoma, breast cancer and medulloblastoma^{47–49}. Moreover members of the JARID1 (also known as KDM5) family that demethylate H3K4me3 and me2 are overexpressed in breast and bladder cancers^{50,51}, and FBXL10 (also known as KDM2B), specific for

H3K36me3 and me2, is overexpressed in leukaemia⁵². Somatic mutations and deletions have also been identified in the JmjC-domain-containing demethylases, including the H3K27me3 and me2 demethylase UTX (also known as KDM6A) that is found mutated in, for instance, multiple myeloma and renal cell carcinoma^{27,28}, and in JARID1C (also known as KDM5C) and PHF8 in patients with X-linked mental retardation^{53,54}. These mutations often lead to loss of a functional demethylase, and because they may be responsible for the disease phenotype, these observations could suggest that the corresponding HMT is a good target for drug development.

Although our understanding of the biological role of the histone demethylases in normal development and disease is still relatively poor, they are considered to be attractive targets for drug development due to their association with disease and their well-defined catalytic mechanism. The use of structure-guided design has recently led to the first highly potent and selective inhibitors to JmjC-domain containing enzymes⁵⁵. These inhibitors, which are competitive with 2-oxoglutarate and non-competitive with a peptide substrate, are potent inhibitors with an half-maximal inhibitory concentration (IC₅₀) in the nanomolar range, and were shown to be specific for the JMJD3 (also known as KDM6B) and UTX H3K27 demethylases. JMJD3 has previously been associated with inflammatory responses, and in agreement with this a JMJD3 and UTX inhibitor reduced proinflammatory cytokine production by human primary macrophages⁵⁵. In addition to showing the relevance of the catalytic activity of JMJD3 in this process, this study provided proof of concept for generating specific JmjC-domain inhibitors. Further proof of concept has been provided by the biotech company EpiTherapeutics, which has developed highly potent inhibitors to the JARID1 family (L.-O. Gerlach, personal communication). These compounds show specific *in vivo* target engagement of JARID1B, an increase in H3K4me3 levels in treated cells and reduced proliferation of cancer cells in a xenograft mouse model (L.-O. Gerlach, personal communication). These proof-of-concept studies provide support for the idea that JmjC-domain-containing proteins can be targeted by specific compounds, which may have therapeutic applications.

LSD1

It is likely that the first small molecule inhibitors of histone demethylases that enter clinical trials will target LSD1 (ref. 56) (Fig. 3c). Several data have suggested that LSD1 could be an interesting therapeutic target in cancer because of its high-level expression in prostate cancer, undifferentiated neuroblastoma, oestrogen-negative breast cancer, bladder cancer and colorectal cancer^{57–60}. Nevertheless, the recent demonstration that LSD1 is required for the development and maintenance of acute myeloid leukaemia (AML) has gained the most attention^{61,62}. Specifically, both genetic and pharmacological data have been provided *in vitro* and in animal models showing that LSD1 is required to sustain the expression of genes induced by the MLL–AF9 oncoprotein and therefore the maintenance of leukaemia stem cells⁶¹. The pharmacological results included the use of the general monooxidase inhibitor tranilcypromine (TCP)⁶² and the TCP-derivative *trans-N*-((2-methoxy-pyridin-3-yl)methyl)-2-phenylcyclopropan-1-amine — developed by the biotech company

Oryzon Genomics⁶³ (Fig. 3c) — that is more specific and 100-fold more potent than TCP⁶¹. The inhibition of LSD1 in AML led to increased differentiation followed by apoptosis, and consistent with this an increase in expression of differentiation markers (for example, CD11b). The inhibition of LSD1 activity was not associated with a global increase in H3K4me2; however, some increase in H3K4me2 was observed on MLL–AF9 bound genes and genes involved in differentiation^{61,62}. Taken together these studies provide proof of concept for LSD1 as a therapeutic target in leukaemia; however, the mechanism by which LSD1 contributes to leukaemia is not clear for several reasons. First, LSD1 has been found to be part of several chromatin complexes, including the neuronal silencer co-repressor of RE1-silencing transcription factor (CoREST; also known as RCOR1) and the nucleosome remodelling and histone deacetylase NuRD⁴⁵ (Fig. 3d). These complexes are found throughout the genome and have a pleiotropic role in transcriptional regulation. Second, LSD1 also binds throughout the genome, especially at active promoters and enhancers^{64,65}. Third, as mentioned above, LSD1 can demethylate H3K9me2 and me1, and H3K4me2 and me1 (Fig. 3d). H3K9me2 is normally found associated with repressed chromatin and transcriptional silencing, whereas H3K4me2 and me1 are associated with active promoters and enhancers. Inhibition of LSD1 activity in AML did not lead to any change in H3K9me2, whereas an increase of H3K4me2 was observed on MLL–AF9 target genes⁶¹ and CD11b⁶². These observations raise several questions. First, if LSD1 is bound throughout the genome, why does the inhibition of LSD1 lead to the selective increase of H3K4me2 on specific promoters? Second, the expression of MLL–AF9 target genes is decreased in response to LSD1 inhibition, whereas H3K4me2 is increased. This is counterintuitive, because an increase in H3K4me2 is normally associated with increased expression of a gene, as

is the case for CD11b. Therefore, what is the mechanism leading to the decreased expression of MLL–AF9 target genes, and how does inhibition of LSD1 lead to differentiation and apoptosis?

Despite the lack of precise mechanistic insight into how LSD1 inhibition can lead to inhibition of leukaemia and prolonged survival of mice, the LSD1 inhibitors seem very promising. Oryzon Genomics has reported on the further development of a clinical compound, ORY-1001, which is more than 1,000 times more potent than TCP and highly selective over related enzymes, including LSD2 (ref. 66). The structure of ORY-1001 has not been revealed; however, it has been shown to reduce leukaemic stem-cell potential, colony formation and to induce differentiation of AML cell lines at subnanomolar concentrations⁶³. Moreover, ORY-1001 leads to the time/dose-dependent increase of H3K4me2 at LSD1 target genes (for example, those that encode CD11b) and induction of differentiation markers (T. Maes, personal communication). Oryzon Genomics expects to take ORY-1001 into phase I clinical trials later this year.

Interestingly, the potential use of LSD1 inhibitors is not limited to oncological disease. In fact, the weak LSD1 inhibitor TCP has been used as a non-selective monoamine oxidase inhibitor for the treatment of depression⁶⁷, and because aberrant activity of the REST–CoREST–LSD complex has been implicated in Huntington's disease⁶⁸ and LSD1 in herpes infection⁶⁹ the LSD1 inhibitors may also be useful for these indications.

Targeting bromodomains

Bromodomains comprise a small family of proteins that recognize and bind to acetylated lysine residues on histone tails (Fig. 4a). Acting as a scaffold for both the assembly of larger, multi-component macromolecular complexes regulating chromatin accessibility and for the recruitment of key transcriptional proteins such as RNA polymerase,

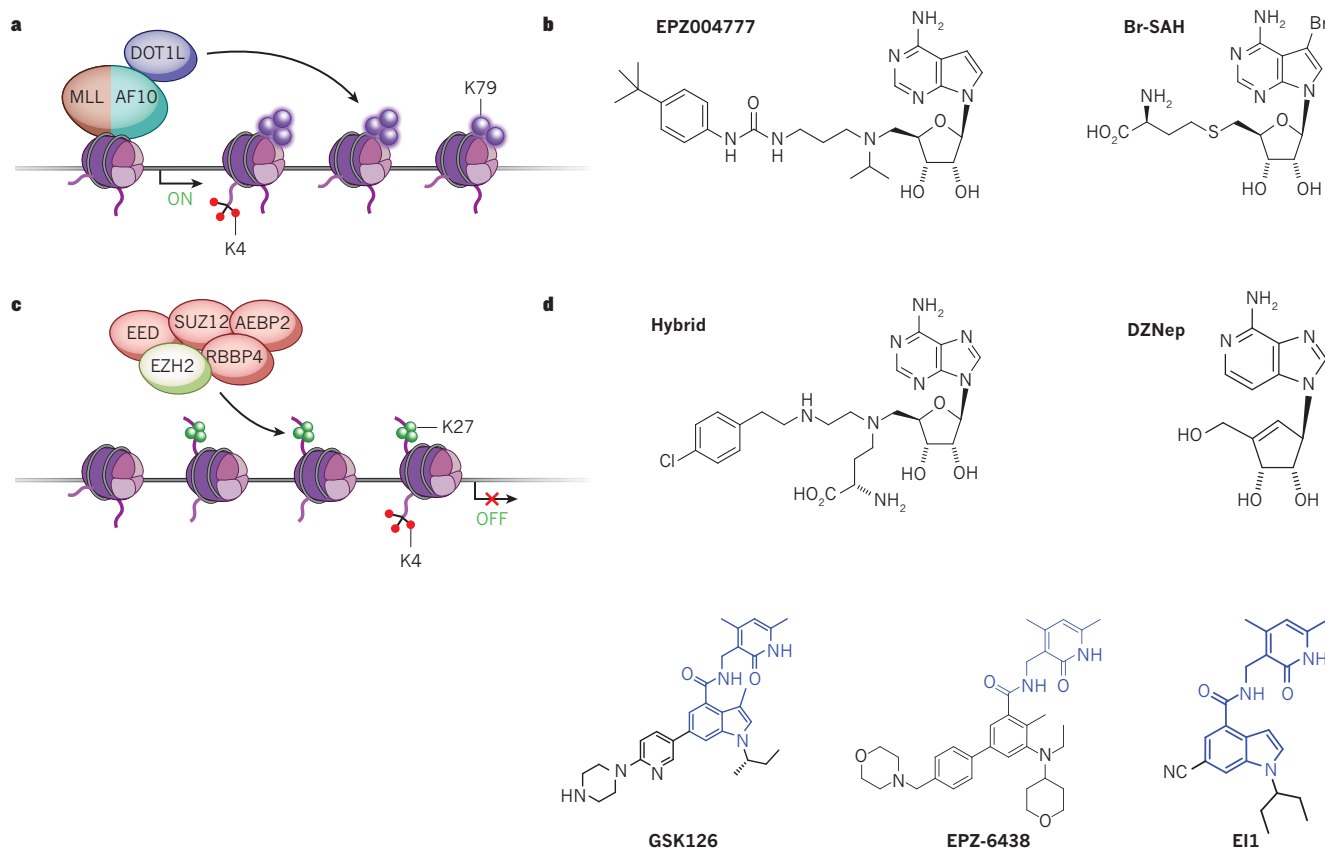


Figure 2 | Histone methyltransferases and inhibitors to DOT1L and EZH2. **a**, DOT1L catalyses H3K79 methylation of nucleosomes associated with actively transcribed genes. It is recruited by MLL-fusion proteins (here exemplified by MLL–AF10) to MLL-target genes, and is required for leukaemia induced by MLL-fusion proteins. **b**, Specific inhibitors to DOT1L are EPZ004777 (ref. 21) and bromo-deaza S-adenosylhomocysteine

(Br-SAH) (ref. 22). **c**, PRC2 catalyses dimethylation and trimethylation of H3K27 to maintain transcription repression of target genes. These target genes are often associated with H3K4me3 as well — a mark of CpG-islands and transcription start sites. **d**, Reported EZH2 inhibitors are a hybrid molecule (ref. 41), 3-deazaneplanocin (DZNep)⁴², GSK126 (ref. 37), EPZ-6438 (ref. 44) and EI1 (ref. 39).

bromodomain-containing proteins are considered 'readers' of the histone code. The human genome encodes more than 50 bromodomain proteins, which can be phylogenetically segregated into eight sub-families⁷⁰. Embryonic lethality on knockdown of the genes encoding bromodomain containing proteins⁷¹ underscores the primary importance of the proteins in basic cell function, but it has also limited our better understanding of their role in normal and disease physiology. Structurally, bromodomains are made up of a bundle of four alpha helices joined by two closely interacting but sequence variable loops that form an invaginated, largely hydrophobic pocket for binding to the acetylated lysine ligand⁷⁰.

The current intense interest in therapeutically targeting various bromodomains originated in the demonstration by GlaxoSmithKline (GSK), the SGC and the Bradner lab that the bromodomain and extra-terminal (BET) subfamily (Brd2, Brd3, Brd4 and BrdT) could be targeted by small molecule antagonists^{72,73}. By directly binding to the BET proteins, such compounds prevent the interaction of the reader module to the acetylated histone thereby preventing assembly of an active gene transcriptional complex (Fig. 4a). The ability to disrupt these protein-protein interactions with drug-like compounds is remarkable and has been shown in multiple structural studies⁶⁷ to be related to the presence of well-defined, deep acetyl lysine binding pockets within the BET proteins. By applying cell-based, high throughput screening of compound libraries combined with elegant chemoproteomics and a battery of structural and biophysical assays, GSK developed compounds that were able to inhibit all four BET proteins but with good selectivity over other bromodomains. Similarly, the SGC working with the Bradner

lab developed the widely used JQ1 (Fig. 4b), originating from a patent application by Mitsubishi–Tanabe⁷⁴. Crucially, the free availability of these compounds to the research community has drastically accelerated our understanding of the primary mechanism of transcriptional regulation and wider chromatin biology. Indeed, the realization that the pharmacological effects of BET inhibition could potentially be applied to ameliorate diverse disease phenotypes has spurred further rounds of compound discovery in pharmaceutical companies.

Early evidence for the potential involvement of BET proteins in cancer was the observation that overexpression of Brd2 in lymphocytes induced B-cell lymphomas. Subsequently, French *et al.* reported that chromosomal translocation of the *Brd4* gene with the NUT protein was the driver for proliferation in the rare but lethal malignancy, NUT-midline carcinoma (NMC)⁷⁵. Furthermore, reversal of the tumour phenotype with BET inhibition not only provided support for the underlying mechanism but also illustrated the therapeutic potential of BET antagonism. Based on this data, a phase I clinical study of the GSK BET inhibitor IBET762 (Fig. 4b) in NMC was initiated in March 2012.

Investigation of the anti-proliferative activity of BET inhibitors in models of haematological cancer, including AML, Burkitt's lymphoma, multiple myeloma and B-cell acute lymphoblastic leukaemia has revealed perhaps the most exciting facet of bromodomain biology^{76,77}. In these malignancies, BET inhibitors such as JQ1 and the more highly bioavailable IBET151 (Fig. 4b) directly silenced MYC expression through disruption of BET protein binding at the MYC locus. Because the various MYC isoforms are known to be crucial

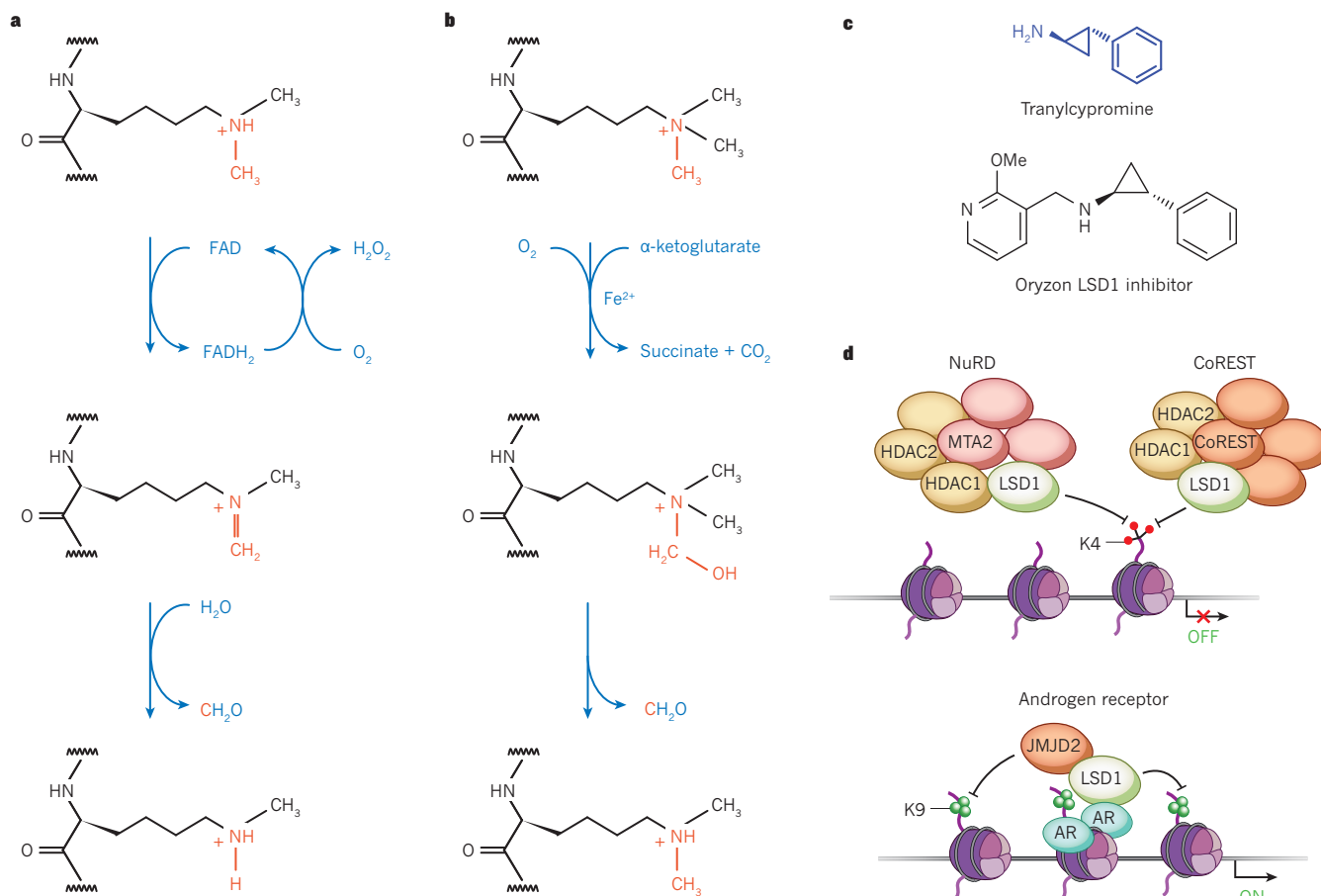


Figure 3 | Histone demethylases and inhibitors to LSD1. **a**, Reaction mechanism used by FAD-dependent LSD1 and LSD2 for demethylation (modified from ref. 45). **b**, Reaction mechanism used by JmjC-domain-containing histone demethylases (modified from ref. 45). **c**, Inhibitors to LSD1. The general monooxidase inhibitor tranylcypromine, and the derivative *trans*-*N*-(2-methoxypyridin-3-yl)methyl)-2-phenylcyclopropan-1-amine)

developed by the biotech company Oryzon Genomics⁶³. **d**, LSD1 is part of several chromatin complexes, including nucleosome remodelling and histone deacetylase (NuRD) and the neuronal silencer CoREST, in which it catalyses the demethylation of H3K4me2 and H3K4me1. As an associated protein with the androgen receptor, together with JMJD2 histone demethylases, LSD1 is responsible for the demethylation of H3K9me2 and H3K9me1 (refs 102,103).

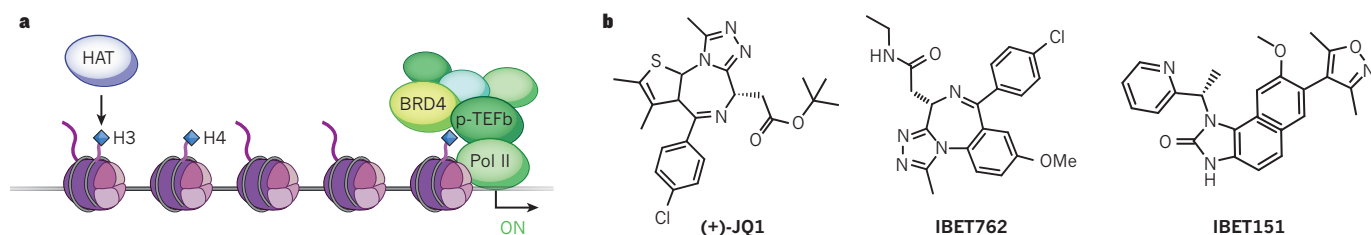


Figure 4 | Bromodomain proteins and their inhibitors. **a**, The bromodomain can bind acetylated lysines, which are associated with actively transcribed promoters. The bromodomain proteins (here illustrated by BRD4 and associated proteins) have a variety of functions, including mediating the initiation and elongation of transcription. BRD4 interacts with positive transcription elongation factor b (p-TEFb), which phosphorylates the C-terminal domain of RNA polymerase II (Pol II) and

induces transcriptional elongation. The interaction of BRD4 with a number of protein complexes involved in transcriptional regulation has also been described. **b**, Chemical structures of prototypical bromodomain and extra-terminal (BET) inhibitors. (+)-JQ1, IBET762 and IBET151 bind to all members of the BET sub-family (Brd2, Brd3, Brd4 and BrdT) with similar affinity and regulate the transcription of key oncogenes including the MYC family and BCL2.

regulators of cell proliferation and survival and MYC is a potent oncogene overexpressed in many cancers, bromodomain antagonism offers, for the first time, an opportunity to target MYC-driven oncogenicity. Intriguingly, however, recent reports have shown crucial subtleties in the mechanism of BET inhibitor modulation of MYC⁷⁸. Whereas in haematological cancers, BET regulates c-MYC, in neuroblastoma, BET inhibitor effects seem to be manifested through silencing of N-MYC, presumably by the same or at least a similar mechanism. These results suggest there is potential for a broader spectrum of activity for BET inhibitors beyond NMC and haematological malignancies, and ongoing clinical studies with IBET762 now include other solid tumours such as N-MYC-amplified lung and colorectal cancers. The question of a therapeutic window for BET inhibitors in a clinical setting remains to be answered but presumably data from animal toxicity studies did not preclude advancing these compounds to human trials.

Outside of cancer, BET inhibition has shown striking effects in a range of inflammatory disease models, suggesting a central role in lymphocyte lineage aetiology. Interestingly, BET inhibition with IBET762 attenuated only secondary response genes in macrophages with no effect on the primary response elements⁷². The ability to modulate selectively the expression of gene subsets is of significance and raises the possibility of further fine-tuning the level of transcriptional activity with selective inhibitors of other bromodomains, which could translate to clinical benefits with fewer undesirable side effects. In mouse models of sepsis, pretreatment with a BET inhibitor suppressed cytokine expression and protected the animals from lethal lipopolysaccharide challenge. In a noteworthy demonstration of activity, administration of the inhibitor even after allergen challenge led to survival⁶⁹. Evidence of the function of other bromodomains (SP110, SPI40 and SMARCA4) in immune-mediated diseases driven by loss of memory T cells and B cells is emerging and limited to tantalizing association of bromodomain expression and disease phenotype. It is too early to say whether small molecule inhibitors of other bromodomains or methyl-lysine readers can be successfully identified, but some promising advances have recently been made with BAZ2B and chromodomain proteins associated with brain tumours (Table 1). The development and availability of additional specific small molecule probes will be needed to help delineate the biology of these proteins.

Perspectives

This is a very exciting and fruitful time for the 'epigenetics field' as illustrated by recent discoveries of new classes of enzymes, insight into the biological role of chromatin-associated proteins, findings showing that somatic mutations in genes coding for chromatin-associated proteins are very frequent in cancer and the development of highly potent and specific small molecule inhibitors to chromatin-associated proteins that show great promise in preclinical trials. Until recently, it was uncertain whether it would be technically feasible to generate specific and potent

inhibitors to the different classes of readers, writers and erasers of the histone code. However, as we have discussed in this Review, this has indeed been possible for very diverse enzymatic classes, such as the HMTs, the two different subclasses of histone demethylases and for the non-enzymatic bromodomain-containing proteins. These inhibitors are undergoing or will shortly enter human phase I clinical trials for a variety of oncology indications albeit initially in rare tumour types or haematopoietic malignancies.

A major challenge for a potential expansion of the inhibitors to other tumour types will be to gain a better understanding of the mechanism of action of the drugs, and therefore of the biology of the target protein. The ongoing phase I clinical trials have all been designed based on genetic evidence for a role of the targeted protein in the disease (DOT1L and LSD1 in AML, EZH2 in DLBCL and IBET in NUT-midline carcinoma). Such strong genetic evidence does not currently exist in other tumour types; however, the effect of the specific inhibitors on large, 'omically' well-characterized cell-line panels will hopefully help to identify specific genetic alterations that lead to drug sensitivity. Nonetheless, even this approach is unlikely to be straightforward because most chromatin-associated proteins are present in several different multi-component complexes that are associated with several thousand genes and loci throughout the genome. The biology is therefore complex and, depending on the tissue and the underlying genetic landscape of the cell, the chromatin-associated protein could act as an oncogene in one setting but be a tumour suppressor in other circumstances. This is illustrated, for instance, by EZH2, in which gain-of-function mutations promote lymphoid transformation^{16,33,37,79} and loss-of-function mutations promote myelodysplastic syndrome and T-cell acute lymphoblastic leukaemia^{36,80–83}. Similarly, somatic mutations of lysine 27 of H3.3 found in paediatric glioblastoma have been shown to inhibit EZH2 activity⁸⁴. The dual roles of EZH2 and H3K27 methylation might also reflect the biological role of EZH2 and the PRC2 complex. In contrast to signalling pathways and transcription factors, chromatin-associated proteins and epigenetic regulation do not seem to be decisive for lineage choice during differentiation. Instead these proteins are present in the genome to ensure transcriptional patterns and cell identity. In other words, the chromatin-associated proteins often fine-tune transcriptional patterns, and the genes regulated by the proteins can be both oncogenes and tumour suppressor genes. The functions of the chromatin-associated proteins do not mean that inhibitors of these proteins will not have a clinical benefit, but highlight the difficulty in identifying biomarkers predictive of tumour sensitivity. This is illustrated again by the EZH2 inhibitors, whereby the levels of EZH2 in a tumour cell line do not predict whether the cell line will respond to the inhibitor; however, a weak correlation does exist between the ability of EZH2 inhibitors to decrease H3K27me3 levels in DLBCL and inhibition of cell growth³⁷.

The generation of small molecule inhibitors of different classes of chromatin-associated proteins has not only increased confidence in the druggability of many epigenetic modulators, but has also provided

strong insights into the rational design of new compounds with higher affinity and specificity. The hope is that this knowledge can be translated into the generation of specific inhibitors of the many other chromatin-associated proteins involved in cancer. At the very least, such inhibitors will be useful as research compounds to understand the biological function of new chromatin-associated proteins, but could eventually also allow for the identification and therapeutic targeting of other pathways that are important for the cancer phenotype. Increasingly, it is becoming evident that effective, long-term responses to anti-cancer therapies require suppression of two or more oncogenic pathways and this is likely to be the case for epigenetic therapies as well. However, modulation of the cancer epigenome with specific inhibitors may offer unique opportunities to discover effective combination therapies based on the potential to directly alter acquired transcriptional resistance mechanisms. Indeed, a recent report⁸⁵ demonstrating reversal of platinum resistance with HDAC inhibition in ovarian cancer highlights such opportunities. Undoubtedly, other rational combinations remain to be identified and the challenge will be to understand the fundamental cellular alterations induced by epigenetic modulators and to develop complementary agents that synergize most effectively. Along these lines, the resurgence and current success of immunotherapeutic approaches to cancer treatment also offers opportunities for epigenetically targeted therapeutics. In principle, it may be possible to induce cell surface expression of tumour-specific antigens, allowing for more effective and sustained immune responses to tumours. Finally, the ability to silence crucial oncogenes such as *MYC* and *BCL2* with bromodomain inhibitors has been remarkable and unpredicted. Inactivation of the master oncogenic proteins with small molecules has been the holy grail for anti-cancer approaches for many years. Yet even here, the lack of a detailed mechanistic understanding of how the BET inhibitors work has led to an empiric approach to determine how best to deploy these agents in the clinic. Despite these limitations, it is important to remember that we are nonetheless on the verge of advancing new molecules with novel biology to human studies with at least some molecular or pathway basis for selecting patients who are most likely to benefit from these agents. Data from these studies will ultimately determine whether these new epigenetic therapies will be a meaningful addition to the armamentarium of physicians, but the signs are promising. ■

Received 31 May; accepted 9 August 2013.

- Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
- Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
- Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nature Rev. Cancer* **11**, 726–734 (2011).
- This review provides an excellent overview of epigenetics with an emphasis on the linkage between genomic and epigenomic phenomena in cancer together with opportunities for biomarker-driven development of therapeutics.**
- Issa, J. P. & Kantarjian, H. M. Targeting DNA methylation. *Clin. Cancer Res.* **15**, 3938–3946 (2009).
- Yang, X., Lay, F., Han, H. & Jones, P. A. Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol. Sci.* **31**, 536–546 (2010).
- Khan, O. & La Thangue, N. B. HDAC inhibitors in cancer biology: emerging mechanisms and clinical applications. *Immunol. Cell Biol.* **90**, 85–94 (2012).
- Wagner, J. M., Hackanson, B., Lubbert, M. & Jung, M. Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin. Epigenetics* **1**, 117–136 (2010).
- You, J. S. & Jones, P. A. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* **22**, 9–20 (2012).
- Huang, J. et al. p53 is regulated by the lysine demethylase LSD1. *Nature* **449**, 105–108 (2007).
- Wu, S. C. & Zhang, Y. Role of protein methylation and demethylation in nuclear hormone signaling. *Mol. Endocrinol.* **23**, 1323–1334 (2009).
- Zhang, K. & Dent, S. Y. Histone modifying enzymes and cancer: going beyond histones. *J. Cell. Biochem.* **96**, 1137–1148 (2005).
- Robinson, G. et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* **488**, 43–48 (2012).
- Ryan, R. J. & Bernstein, B. E. Genetic events that shape the cancer epigenome. *Science* **336**, 1513–1514 (2012).
- Shih, A. H., Abdel-Wahab, O., Patel, J. P. & Levine, R. L. The role of mutations in epigenetic regulators in myeloid malignancies. *Nature Rev. Cancer* **12**, 599–612 (2012).

- McCabe, M. T. et al. Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypermethylation of histone H3 on lysine 27 (H3K27). *Proc. Natl Acad. Sci. USA* **109**, 2989–2994 (2012).
- Sneering, C. J. et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypermethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl Acad. Sci. USA* **107**, 20980–20985 (2010).
- Kuo, A. J. et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* **44**, 609–620 (2011).
- Meyer, C. et al. New insights to the MLL recombinome of acute leukemias. *Leukemia* **23**, 1490–1499 (2009).
- Okada, Y. et al. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
- Bernt, K. M. et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).
- Daigle, S. R. et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53–65 (2011).
- References 20 and 21 present compelling evidence for the involvement of DOT1L and H3K79 methylation in MLL-rearranged leukaemia and provide rationale for therapeutic targeting of the enzyme.**
- Anglin, J. L. et al. Synthesis and structure-activity relationship investigation of adenosine-containing inhibitors of histone methyltransferase DOT1L. *J. Med. Chem.* **55**, 8066–8074 (2012).
- Yu, W. et al. Bromo-deaza-SAH: a potent and selective DOT1L inhibitor. *Bioorg. Med. Chem.* **21**, 1787–1794 (2013).
- Daigle, S. R. et al. Potent inhibition of DOT1L as treatment for MLL-fusion leukemia. *Blood* <http://dx.doi.org/10.1182/blood-2013-04-497644> (2013).
- Morey, L. & Helin, K. Polycomb group protein-mediated repression of transcription. *Trends Biochem. Sci.* **35**, 323–332 (2010).
- Varambally, S. et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
- Dalglish, G. L. et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
- van Haften, G. et al. Somatic mutations of the histone H3K27 demethylase gene *UTX* in human cancer. *Nature Genet.* **41**, 521–523 (2009).
- Kleer, C. G. et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl Acad. Sci. USA* **100**, 11606–11611 (2003).
- Wagner, N. et al. Enhancer of zeste homolog 2 (EZH2) expression is an independent prognostic factor in renal cell carcinoma. *BMC Cancer* **10**, 524 (2010).
- Takawa, M. et al. Validation of the histone methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a prognostic marker. *Cancer Sci.* **102**, 1298–1305 (2011).
- Bracken, A. P. et al. EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J.* **22**, 5323–5335 (2003).
- Morin, R. D. et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genet.* **42**, 181–185 (2010).
- Pasqualucci, L. et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nature Genet.* **43**, 830–837 (2011).
- Ryan, R. J. et al. EZH2 codon 641 mutations are common in BCL2-rearranged germinal center B cell lymphomas. *PLoS ONE* **6**, e28585 (2011).
- Nikoloski, G. et al. Somatic mutations of the histone methyltransferase gene *EZH2* in myelodysplastic syndromes. *Nature Genet.* **42**, 665–667 (2010).
- McCabe, M. T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112 (2012).
- This is the first report of a potent and selective inhibitor of EZH2 with *in vitro* and *in vivo* activity compared with EZH2 mutant lymphomas.**
- Knutson, S. K. et al. A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nature Chem. Biol.* **8**, 890–896 (2012).
- Qi, W. et al. Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc. Natl Acad. Sci. USA* **109**, 21360–21365 (2012).
- Valente, S. et al. Identification of PR-SET7 and EZH2 selective inhibitors inducing cell death in human leukemia U937 cells. *Biochimie* **94**, 2308–2313 (2012).
- Epizyme Inc. Inhibitors of human EZH2, and methods to use thereof. WO/2012/034132 (2012).
- Tan, J. et al. Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev.* **21**, 1050–1063 (2007).
- Miranda, T. B. et al. DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. *Mol. Cancer Ther.* **8**, 1579–1588 (2009).
- Knutson, S. K. et al. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proc. Natl Acad. Sci. USA* **110**, 7922–7927 (2013).
- This article reports activity of EZH2 inhibitors in solid tumours suggesting potential for clinical benefit beyond haematological malignancies.**
- Kooistra, S. M. & Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nature Rev. Mol. Cell Biol.* **13**, 297–311 (2012).
- Mosammaparast, N. & Shi, Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* **79**, 155–179 (2010).
- Yang, Z. Q. et al. Identification of a novel gene, *GASC1*, within an amplicon at 9p23–24 frequently detected in esophageal cancer cell lines. *Cancer Res.* **60**, 4735–4739 (2000).

48. Liu, G. *et al.* Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer. *Oncogene* **28**, 4491–4500 (2009).
49. Ehrbrecht, A. *et al.* Comprehensive genomic analysis of desmoplastic medulloblastomas: identification of novel amplified genes and separate evaluation of the different histological components. *J. Pathol.* **208**, 554–563 (2006).
50. Lu, P. J. *et al.* A novel gene (*PLU-1*) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. *J. Biol. Chem.* **274**, 15633–15645 (1999).
51. Hayami, S. *et al.* Overexpression of the JmJc histone demethylase KDM5B in human carcinogenesis: involvement in the proliferation of cancer cells through the E2F/RB pathway. *Mol. Cancer* **9**, 59 (2010).
52. He, J., Nguyen, A. T. & Zhang, Y. KDM2b/JHDM1b, an H3K36me2-specific demethylase, is required for initiation and maintenance of acute myeloid leukemia. *Blood* **117**, 3869–3880 (2011).
53. Jensen, L. R. *et al.* Mutations in the *JARID1C* gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *Am. J. Hum. Genet.* **76**, 227–236 (2005).
54. Laumonnier, F. *et al.* Mutations in PPHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *J. Med. Genet.* **42**, 780–786 (2005).
55. Kruidenier, L. *et al.* A selective jumoni H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408 (2012).
This article describes the discovery, structural biology and activity of potent and selective Jumoni demethylase inhibitors.
56. Lynch, J. T., Harris, W. J. & Somervaille, T. C. LSD1 inhibition: a therapeutic strategy in cancer? *Expert Opin. Ther. Targets* **16**, 1239–1249 (2012).
57. Hayami, S. *et al.* Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int. J. Cancer* **128**, 574–586 (2011).
58. Kauffman, E. C. *et al.* Role of androgen receptor and associated lysine-demethylase coregulators, LSD1 and JMJD2A, in localized and advanced human bladder cancer. *Mol. Carcinog.* **50**, 931–944 (2011).
59. Kahl, P. *et al.* Androgen receptor coactivators lysine-specific histone demethylase 1 and four and a half LIM domain protein 2 predict risk of prostate cancer recurrence. *Cancer Res.* **66**, 11341–11347 (2006).
60. Schulte, J. H. *et al.* Lysine-specific demethylase 1 is strongly expressed in poorly differentiated neuroblastoma: implications for therapy. *Cancer Res.* **69**, 2065–2071 (2009).
61. Harris, W. J. *et al.* The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21**, 473–487 (2012).
62. Schenk, T. *et al.* Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-*trans*-retinoic acid differentiation pathway in acute myeloid leukemia. *Nature Med.* **18**, 605–611 (2012).
Reference 61 and 62 highlight the role of LSD1 in AML and the potential for inhibitors to synergize with all-*trans*-retinoic acid therapy.
63. Oryzon Genomics. Phenylcyclopropylamine derivatives and their medical use. WO/2010/084160 (2010).
64. Ram, O. *et al.* Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* **147**, 1628–1639 (2011).
65. Whyte, W. A. *et al.* Enhancer decommisioning by LSD1 during embryonic stem cell differentiation. *Nature* **482**, 221–225 (2012).
66. Maes, T. *et al.* Preclinical characterization of a potent and selective inhibitor of the histone demethylase KDM1A for MLL leukemia. *J. Clin. Oncol.* **31**, suppl; abstr. e13543 (2013).
67. Lee, M. G., Wynder, C., Schmidt, D. M., McCafferty, D. G. & Shiekhkhattar, R. Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. *Chem. Biol.* **13**, 563–567 (2006).
68. Zuccato, C. *et al.* Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nature Genet.* **35**, 76–83 (2003).
69. Liang, Y., Vogel, J. L., Narayanan, A., Peng, H. & Kristie, T. M. Inhibition of the histone demethylase LSD1 blocks α -herpesvirus lytic replication and reactivation from latency. *Nature Med.* **15**, 1312–1317 (2009).
70. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214–231 (2012).
71. Gyuris, A. *et al.* The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. *Biochim. Biophys. Acta* **1789**, 413–421 (2009).
72. Nicodeme, E. *et al.* Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119–1123 (2010).
73. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
74. Mitsubishi-Tanabe Pharma Corporation. Antitumor agent. WO/2009/084693 (2009).
75. French, C. A. Demystified molecular pathology of NUT midline carcinomas. *J. Clin. Pathol.* **63**, 492–496 (2010).
76. Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).
77. Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904–917 (2011).
Together with references 72 and 73, references 76 and 77 are outstanding demonstrations of the feasibility of inhibiting bromodomain proteins in inflammation and tumorigenesis.
78. Puissant, A. *et al.* Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
79. Béguelin, W. *et al.* EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* **23**, 677–692 (2013).
80. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nature Genet.* **42**, 722–726 (2010).
81. Ntziachristos, P. *et al.* Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nature Med.* **18**, 298–301 (2012).
82. Simon, C. *et al.* A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes Dev.* **26**, 651–656 (2012).
83. Zhang, J. *et al.* The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157–163 (2012).
84. Lewis, P. W. *et al.* Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340**, 857–861 (2013).
85. Stronach, E. A. *et al.* HDAC4-regulated STAT1 activation mediates platinum resistance in ovarian cancer. *Cancer Res.* **71**, 4412–4422 (2011).
86. Yu, W. *et al.* Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nature Commun.* **3**, 1288 (2012).
87. Amatangelo, M. D. *et al.* Three-dimensional culture sensitizes epithelial ovarian cancer cells to EZH2 methyltransferase inhibition. *Cell Cycle* **12**, 2113–2119 (2013).
88. Structural Genomics Consortium. *Chemical Probes*. http://www.thesgc.org/scientists/chemical_probes (SGC, 2013).
89. Konze, K. D. *et al.* An orally bioavailable chemical probe of the lysine methyltransferases EZH2 and EZH1. *ACS Chem. Biol.* **8**, 1324–1334 (2013).
90. Kubicek, S. *et al.* Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol. Cell* **25**, 473–481 (2007).
91. Liu, F. *et al.* Protein lysine methyltransferase G9a inhibitors: design, synthesis, and structure activity relationships of 2,4-diamino-7-aminoalkoxy-quinazolines. *J. Med. Chem.* **53**, 5844–5857 (2010).
92. Vedadi, M. *et al.* A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nature Chem. Biol.* **7**, 566–574 (2011).
93. Yuan, Y. *et al.* A small-molecule probe of the histone methyltransferase G9a induces cellular senescence in pancreatic adenocarcinoma. *ACS Chem. Biol.* **7**, 1152–1157 (2012).
94. Liu, F. *et al.* Exploiting an allosteric binding site of PRMT3 yields potent and selective inhibitors. *J. Med. Chem.* **56**, 2110–2124 (2013).
95. Wan, H. *et al.* Benzo[d]imidazole inhibitors of co-activator associated arginine methyltransferase 1 (CARM1) — hit to lead studies. *Bioorg. Med. Chem. Lett.* **19**, 5063–5066 (2009).
96. Sack, J. S. *et al.* Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem. J.* **436**, 331–339 (2011).
97. Allan, M. *et al.* N-Benzyl-1-heteroaryl-3-(trifluoromethyl)-1H-pyrazole-5-carboxamides as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **19**, 1218–1223 (2009).
98. Seal, J. *et al.* Identification of a novel series of BET family bromodomain inhibitors: binding mode and profile of I-BET151 (GSK1210151A). *Bioorg. Med. Chem. Lett.* **22**, 2968–2972 (2012).
99. Picaud, S. *et al.* PFI-1, a highly selective protein interaction inhibitor, targeting BET bromodomains. *Cancer Res.* **73**, 3336–3346 (2013).
100. Herold, J. M. *et al.* Small-molecule ligands of methyl-lysine binding proteins. *J. Med. Chem.* **54**, 2504–2511 (2011).
101. James, L. I. *et al.* Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nature Chem. Biol.* **9**, 184–191 (2013).
102. Metzger, E. *et al.* LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437**, 436–439 (2005).
103. Wissmann, M. *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nature Cell Biol.* **9**, 347–353 (2007).

Acknowledgement Work in the Helin laboratory is supported by the Danish National Research Foundation, the Danish Cancer Society, the Novo Nordisk Foundation, the Lundbeck Foundation, the European Union, The European Research Council and the Excellence Programme of the University of Copenhagen.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of this paper at go.nature.com/4tc9uy. Readers are welcome to comment on the online version of this article at go.nature.com/4tc9uy. Correspondence should be addressed to K.H. (kristian.helin@bric.ku.dk) or D.D. (ddhanak8@ITS.JNJ.com).

The nexus of chromatin regulation and intermediary metabolism

Philipp Gut¹ & Eric Verdin¹

Living organisms and individual cells continuously adapt to changes in their environment. Those changes are particularly sensitive to fluctuations in the availability of energy substrates. The cellular transcriptional machinery and its chromatin-associated proteins integrate environmental inputs to mediate homeostatic responses through gene regulation. Numerous connections between products of intermediary metabolism and chromatin proteins have recently been identified. Chromatin modifications that occur in response to metabolic signals are dynamic or stable and might even be inherited transgenerationally. These emerging concepts have biological relevance to tissue homeostasis, disease and ageing.

Cellular phenotypic plasticity — the extent to which a cell can adapt to changes in its environment — is the ultimate determinant of its sustained function and survival¹. Cellular plasticity relies on the precise coordination of transcriptional programs that allow the cell to adapt in the face of a changing environment. A complex set of cellular regulatory mechanisms determine which genes are activated by transcription factors at a given time and in a specific cellular context. The packaging of DNA and histones into chromatin is an important aspect of gene regulation, allowing the access of transcription complexes to DNA to be regulated and for histones to participate in the regulatory process^{2–5}. The smallest building block of chromatin is the nucleosome, which consists of 147 nucleotides of DNA wrapped around an octamer of core histones (two molecules of each of the four histones: H2A, H2B, H3 and H4)⁶. Post-translational modifications (PTMs) of histones and modification to DNA itself in the form of methylation can alter the structure of chromatin and help to recruit transcription factors and other gene regulatory proteins to the DNA^{7,8}. Thus, chromatin-associated modifications modulate the interaction of transcriptional complexes and DNA, thereby influencing the transcriptional network that ultimately regulates gene expression. Importantly, chromatin modifications can be highly dynamic or in a more stable configuration that persists as a ‘transcriptional memory’ through mitosis or meiosis^{2–5,9}.

Strikingly, almost all chromatin-modifying enzymes utilize co-factors or substrates that are crucial metabolites in core pathways of intermediary metabolism. These metabolites include acetyl-CoA, uridine diphosphate (UDP)-glucose, α -ketoglutarate (α -KG), nicotinamide adenine dinucleotide (NAD⁺), flavin adenine dinucleotide (FAD), ATP or S-adenosylmethionine (SAM) (Fig. 1). Because the cellular concentrations of several of these metabolites fluctuate as a function of the metabolic status of the cell, the activity of the chromatin regulators may change as a function of metabolic status and thereby transduce a homeostatic transcriptional response. A compelling body of evidence has accumulated in recent years in support of this hypothesis^{10–14}.

Here, we review this emerging model — that chromatin-associated enzymes sense intermediary metabolism products and process this information into dynamic chromatin PTMs. These chromatin modifications, in turn, help to coordinate homeostatic or adaptive transcriptional responses. In other cases, sensing of metabolic signals can also drive the activity of gene networks that control fundamental cell fate decisions, including pluripotency of stem cells and cancer transformation. Furthermore, we discuss the exciting perspective that disturbances

in energy metabolism might also lead to stable epigenetic changes that are maintained through the germ line and may affect the health of the next generations. The conventional and most restrictive definition of epigenetics is mitotically or meiotically heritable changes in gene function that are independent of any change in DNA sequence. A more recent and broader definition of epigenetics is “a structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states”¹⁵. In this Review, we adhere to the more traditional definition of epigenetics.

Transcriptional links with intermediary metabolism

DNA and histone modifications have a major influence on the control of gene transcription during embryonic development as well as in the differentiated tissues of the adult organism^{16,17}. In particular, PTMs of histone proteins are emerging as a dynamic mechanism to rapidly alter DNA–chromatin interactions in response to intracellular signals, thereby modulating gene regulation. PTMs are covalent, reversible chemical modifications of amino acid residues within a protein that change the function of the protein, its stability, its subcellular localization or its interaction with other proteins. The best-studied PTMs are phosphorylation of threonine or serine residues, as well as ubiquitination and acetylation of lysine residues. In the case of acetylation, a picture has recently emerged that places the PTM of enzymes of intermediary metabolism as a key regulatory mechanism that rapidly adapts metabolite flux to changes in a cell’s energetic state^{18–20}. This Review focuses on the mounting evidence for an analogous concept in which metabolites influence the status of PTMs of histones, thereby dynamically adapting transcriptional programs to metabolic substrate availability.

Post-translational modifications of histones

PTMs of histones occur mostly within the amino-terminal histone ‘tails’ that protrude from their surface³, but also within the histone globular domain. Methylation occupies lysine or arginine residues, acetylation occupies lysine residues and phosphorylation occurs on serine or threonine residues, the latter more frequently within the histone globular domain (except for serine 10 on histone H3). The distribution of histone marks across genes and their regulatory regions has the highest density within the upstream region, the core promoter and the 5′ and 3′ areas of the coding sequence².

Specific chromatin marks, such as methylation of histone 3 lysine 27 (H3K27), are generally associated with repression of gene transcription. Other modifications, such as histone acetylation or trimethylation

¹Gladstone Institutes, University of California, San Francisco, California 94941, USA.

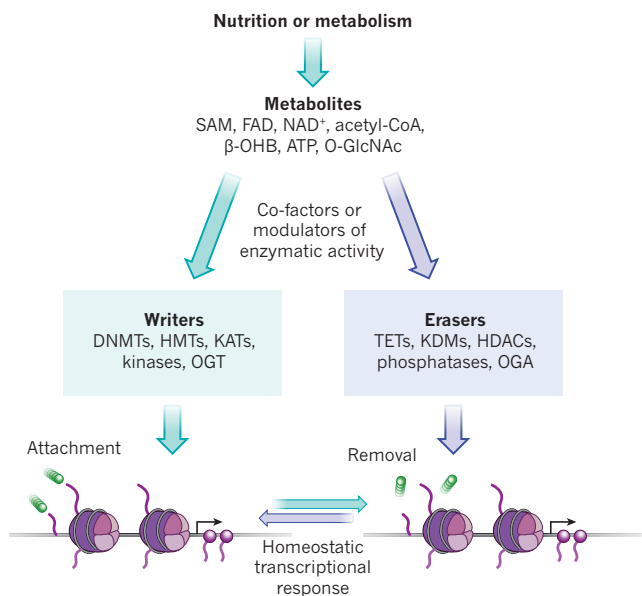


Figure 1 | DNA methylation and post-translational modifications of histones link metabolites and transcription. Changes in nutrition or fluctuations in metabolism induce homeostatic transcriptional responses. Several intermediary metabolism products change enzymatic activity of chromatin-associated proteins in a dose-dependent manner. ‘Writer’ enzymes that attach marks covalently to chromatin or DNA and ‘erasers’ that remove these modifications act as metabolic sensors. Chromatin modifications remodel DNA–histone interactions and help to regulate the recruitment of transcriptional complexes to genes that control cellular function and survival. DNMT, DNA methyltransferases; FAD, flavin adenine dinucleotide; HDACs, histone deacetylases; HMTs, histone methyltransferases; KATs, lysine acetyltransferases; KDMs, lysine demethylases; O-GlcNAc, O-linked *N*-acetylglucosamine; OGT, O-GlcNAc transferase; OGA, O-GlcNAcase; β-OHB, β-hydroxybutyrate; SAM, S-adenosylmethionine; TET, ten-eleven translocation protein.

of histone 3 lysine 4 (H3K4me3), are frequently associated with active transcription^{4,9,21}. However, the consequences of unique modifications are often more complex and depend on the context of all the modifications associated with a given gene locus. This combinatorial readout of epigenetic information has been termed the chromatin language or histone code^{3–5}. Continuing the language analogy, the gene regulatory enzymes that add unique modifications to histones or DNA are referred to as ‘writers’; the proteins that recognize unique PTMs are referred to as ‘readers’; and the enzymes that remove the modifications are referred to as ‘erasers’.

Histone modifications can regulate transcription by changing the biophysical properties of chromatin fibres. For example, acetylation is thought to lead to a decreased interaction between distinct chromatin fibres and to a decondensation of chromatin and increased accessibility of DNA to the transcriptional machinery. The modification of histones is a reversible and tightly regulated biological process. For each PTM, selective enzymes deposit unique marks (Fig. 1). Histone modifications may also facilitate the recruitment of selective readers that recognize single or combinations of histone modifications⁸. The sequential docking to chromatin marks, in turn, helps to recruit proteins that mediate DNA-associated activities, such as transcription, replication and repair⁸.

As introduced earlier, histone-modifying enzymes utilize intermediary metabolism products as substrates or co-factors. For example histone acetylation by lysine acetyltransferases (KATs) depends on intracellular levels of acetyl-CoA, often referred to as ‘activated acetate’, for its high energetic state. The inverse reaction, removal of acetyl groups from histones, is mediated by histone deacetylases (HDACs). The class III histone deacetylases, sirtuin proteins, consume the energy carrier NAD⁺ as a co-factor^{10,11,13}. Histone methyltransferases (HMTs) and histone demethylases also require metabolites for their enzymatic activity. HMTs use the methyl group donor SAM¹⁴. The first histone demethylase discovered,

LSD1, removes methyl groups from histones in a flavin-dependent oxidative reaction²² and the histone demethylases of the Jumoni class (JM) rely on Fe(II) and α-KG for their enzymatic activity²³.

Histone glycosylation through O-linked *N*-acetylglucosamine (O-GlcNAc) modification of serine and threonine is a recently recognized modification²⁴. In the roundworm *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, mice and humans, addition and removal of O-GlcNAc are catalysed by two enzymes, O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA), respectively²⁵. Because UDP-glucose (the substrate for O-GlcNAcylation) is a product of the hexosamine pathway (an alternative metabolic branch of glucose metabolism) and directly reflects changes in ambient glucose levels, this modification probably links intermediary metabolism with a unique chromatin modification²⁶. All four core histone proteins can be glycosylated by OGT on sites that can alternatively be phosphorylated^{24,27}. In addition, non-histone proteins in chromatin, such as members of the Polycomb and Trithorax group complexes in *D. melanogaster*, can also be modified through O-GlcNAcylation²⁶.

Although this Review focuses on chromatin, it should be noted that chromatin-modifying enzymes also modify non-histone proteins. For example, the NAD⁺-dependent deacetylase SIRT1 not only deacetylates histones, but also non-histone targets that are crucial for cellular energy metabolism, such as the transcriptional co-activator PGC-1α²⁸. Both histone and non-histone protein modifications are therefore likely to be influenced by changes in metabolite concentrations and both are probably important for the biological activities of the enzymes.

DNA methylation

Methylation of cytosine (at the carbon-5 position) at CpG dinucleotides is the predominant epigenetic modification of DNA in vertebrates²⁹. DNA methylation affects gene activity directly by inhibiting the binding of transcription factors and indirectly by recruiting chromatin-associated proteins with repressive properties⁷. DNA methylation contributes to cell-lineage restriction and genetic imprinting, such as the silencing of the X chromosome in female mammals, during development. DNA methylation is catalysed either by maintenance methyltransferases (such as DNMT1) that add methyl groups to hemimethylated DNA during replication, or by *de novo* methyltransferases (DNMT3a and DNMT3b) that are active after completion of replication¹⁷. The removal of DNA methylation is regulated by TET proteins, which convert 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine. These modifications can be removed during DNA replication or by iterative oxidation and base excision repair. The mechanism of DNA demethylation by TET proteins and its biological role is reviewed by Kohli and Zhang³⁰ in this Insight.

Although DNA methylation is considered relatively stable, dynamic changes occur in adult tissues, and at least some of these adaptations may occur in response to metabolic stimuli^{1,31,32}. Interestingly, the genes encoding PGC-1α, pyruvate dehydrogenase kinase 4 (PDK4), and mitochondrial transcription factor A (TFAM) show hypermethylated DNA stretches within their promoter sequences in skeletal muscle of patients with type 2 diabetes³³ and DNA methylation decreases at these loci in response to an acute bout of exercise³². More recent data suggest that normalization of methylation status of *PGC-1α* and other metabolic genes occurs after gastric bypass surgery, an effective intervention to reduce weight in humans who are obese³⁴. Thus, although the mechanisms that govern dynamic changes in DNA methylation in post-mitotic tissues remain largely obscure, this epigenetic modification probably contributes to homeostatic transcriptional adaptations.

DNA methylation is linked to intermediary metabolism through SAM as a methyl donor substrate. SAM is generated in a cyclic pathway, termed one-carbon metabolism, from the amino acid methionine and ATP. When a methyl group of SAM is transferred to a macromolecule (DNA or histone, see later), the product S-adenosylhomocysteine (SAH) is recycled back to SAM. Tetrahydrofolate, a derivative of folate (vitamin B9) serves as a methyl group donor and combines with glycine to generate methionine, the immediate precursor to SAM. The dependence of

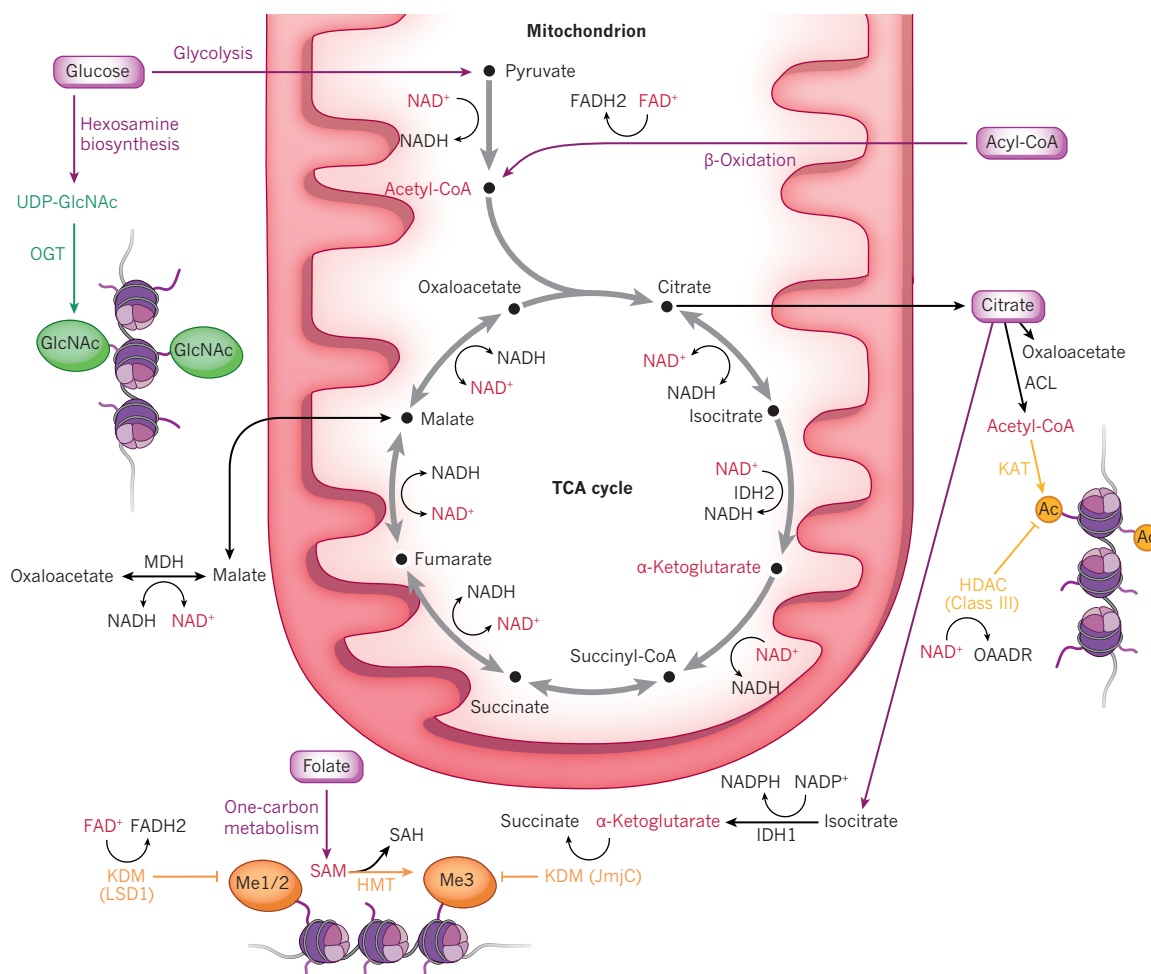


Figure 2 | Metabolic pathways of intermediary metabolism signal to chromatin. Metabolites that regulate chromatin participate in major biochemical pathways involved in intracellular energy balance. The tricarboxylic acid (TCA) cycle is the central hub that links catabolic and anabolic pathways. Glycolysis and β -oxidation are catabolic reactions that generate acetyl-CoA, whereas removal of acetyl-CoA during glucose excess from the mitochondria by the citrate shuttle fuels lipogenesis and biosynthesis of various other macromolecules and together with KAT leads to histone acetylation (shown in yellow). The hexosamine biosynthetic pathway is an alternative route of glucose utilization that generates the co-enzyme UDP-GlcNAc, which together with OGT leads to histone O-GlcNAcylation (shown in green). Folate

(vitamin B9) is a micronutrient that enters one-carbon metabolism, a cyclic reaction generating SAM, as a methyl donor for DNA and histone methylation (shown in orange). Mitochondrial $\text{NAD}^+:\text{NADH}$ ratios are connected with the nuclear cytosolic compartment by the malate–aspartate shuttle. Metabolic regulators of chromatin-modifying enzymes are shown in red. ACL, ATP-citrate lyase; FAD, flavin adenine dinucleotide; GlcNAc, *N*-acetylglucosamine; HDAC, histone deacetylase; HMT, histone methyltransferase; IDH, isocitrate dehydrogenase; KAT, lysine acetyltransferase; KDM, lysine demethylase; MDH, malate dehydrogenase; OAADR, O-acetyl ADP-ribose; OGT, O-linked *N*-acetylglucosamine transferase; SAM, *S*-adenosylmethionine; UDP, uridine diphosphate.

one-carbon metabolism on folate and other micronutrients suggests a direct connection between nutrition and DNA methylation³⁵. Whether these nutrients confer relevant effects on DNA methylation in humans is still under debate, although multiple experimental models point in this direction (discussed later). Removal of 5-methylcytosine by TET proteins depends on Fe(II) and α -KG, an intermediate of the tricarboxylic acid (TCA) cycle and catabolic metabolism of the amino acid glutamine. It is not clear whether TET enzymes sense α -KG levels. However, two other TCA cycle intermediates, fumarate and succinate, inhibit TET, suggesting that the relative concentrations of these metabolites may regulate TET enzymatic activity³⁶.

Sensing of metabolites by chromatin

The idea that gene transcription is influenced by intermediary metabolism products through epigenetic mechanisms was suggested several years ago^{37,38}. Despite the apparent connection between chromatin regulators and their metabolic substrates, the biological relevance of this attractive concept remained largely unexplored until recently. In particular, it wasn't clear whether the chromatin-modifying enzyme co-factors were indeed rate limiting, and whether fluctuations of their local concentrations were

sufficiently dynamic and occurred in concentrations likely to affect the enzymatic activity of their cognate enzymes^{10–14,39,40}. At the centre of this debate is whether chromatin-regulator function is more similar to metabolic enzymes, whose activity depends on the relative abundance of their substrates and products, or more similar to protein kinases, whose activity is relatively independent of physiological fluctuations in ATP levels¹⁰. Although further work is required, recent findings suggest that chromatin-regulating proteins indeed sense intracellular cofactor levels. As discussed later, this phenomenon is particularly well explored for acetylation.

Histone acetylation links metabolism and transcription

In eukaryotes, acetyl-CoA is the universal donor for acetylation reactions and is generated when ingested nutrients enter catabolic pathways of intermediary metabolism (Fig. 2). During fasting, after glycogen stores have been depleted, the organism switches to fatty-acid oxidation, and the acetyl-CoA generated is used to fuel the TCA cycle and thereby oxidative phosphorylation and ATP synthesis. By contrast, during feeding, glycolysis is the primary pathway for acetyl-CoA generation and rapidly satisfies the basic energetic need of a cell. As a result, the excess acetyl-CoA is exported from the mitochondria to the cytosol as citrate. There,

the citrate is re-converted to acetyl-CoA by ATP-citrate lyase and serves as a carbon donor for anabolic reactions, including lipogenesis, and cholesterol and amino acid synthesis¹². Thus, catabolic or anabolic flux leads to changes in free acetyl-CoA levels, a phenomenon that can be observed during the physiological feeding-to-fasting transition in mice. Indeed, fasting is associated with a marked increase in lysine acetylation in hundreds of proteins^{41,42}. An intriguing question is whether the abundance of acetyl-CoA during different metabolic states serves directly as a signal that regulates transcriptional programs.

Landmark studies in yeast show that cytosolic and nuclear acetyl-CoA levels are indeed a crucial determinant of histone acetylation. Unicellular organisms such as yeast base their decision on whether to grow and divide on the environmental abundance of nutrients. When carbon sources are sparse, yeast refrain from entering cellular division to preserve substrates for vital housekeeping processes. On re-exposure to energy sources, cells readily shift back to the coordinated expression of proliferation and growth-related gene clusters. An elegant system to study the relationship between nutrient intake and molecular mechanisms of cell division is the yeast metabolic cycle⁴³. When exposed to a limited nutrient supply, yeast require a 4–5-hour cycle of highly oxidative and reductive metabolism before they can divide. The purpose of cycling through three distinct metabolic stages is to stockpile sufficient building blocks for new cells (oxidative phase) before enough material is acquired to execute division (reductive, building phase, which is followed by the quiescent reductive, charging phase)⁴⁴. Importantly, acetyl-CoA pools fluctuate drastically, about 10-fold, between the oxidative and the reductive phases, and the fluctuations in acetyl-CoA concentrations are within a range likely to affect the enzymatic activity of several KATs⁴⁵. Synchronous to the yeast metabolic cycle-phase oscillations, phase-dependent expression of gene clusters that either promote growth (oxidative growth) or restrict growth (reductive growth) is initiated⁴⁶. Strikingly, the expression of the growth-phase

gene group at the end of the oxidative phase depends on acetyl-CoA. Supplementation of yeast with diverse carbon sources, including glucose, galactose, ethanol or acetate, leads to shortening of the yeast metabolic cycle⁴⁶. To generate acetyl-CoA in the absence of the direct acetyl-CoA sources glucose or galactose, ethanol is converted to acetate, which in turn is activated by acetyl-CoA synthetase. The elevated levels of acetyl-CoA from any of these precursors correlate with histone hyperacetylation and increased expression of the growth-related gene network, including genes that encode ribosomal RNA and protein translation machineries as well as those for amino acid synthesis⁴⁶. Deficiency of the KAT complex SAGA (also known as Gcn5) abrogates both the bulk hyperacetylation of H3 and H4 histones and the expression of growth-related genes. The nuclear cytosolic acetyl-CoA synthetase (Acs2p) is necessary for histone acetylation when ethanol or acetate are the main carbon source⁴⁷. In conclusion, this elegant work in yeast suggests that acetyl-CoA is a driver metabolite of histone acetylation and transcriptional control.

Similarly, observations in mammalian cell culture experiments suggest that histone acetylation also depends on intracellular acetyl-CoA pools⁴⁸. Feeding cells in high glucose concentrations leads to increased glycolysis, pyruvate generation and an increase in mitochondrial acetyl-CoA. Deletion of ATP-citrate lyase suppresses conversion of citrate to nuclear or cytosolic acetyl-CoA and decreases histone acetylation. Although deletion of *AceCS1*, the mammalian homologue of *Acs2p*, does not affect histone acetylation, high levels of acetate (5 mM) in the culture medium can substitute for ATP-citrate lyase deficiency. These results suggest that histone acetylation primarily depends on glucose-derived, cytosolic pools of acetyl-CoA. Histone acetylation during high-energy conditions induces a transcriptionally permissive chromatin configuration that allows a feed-forward control mechanism for the selective expression of genes that regulate cellular proliferation, lipogenesis and adipocyte differentiation^{12,48} (Fig. 3a). In

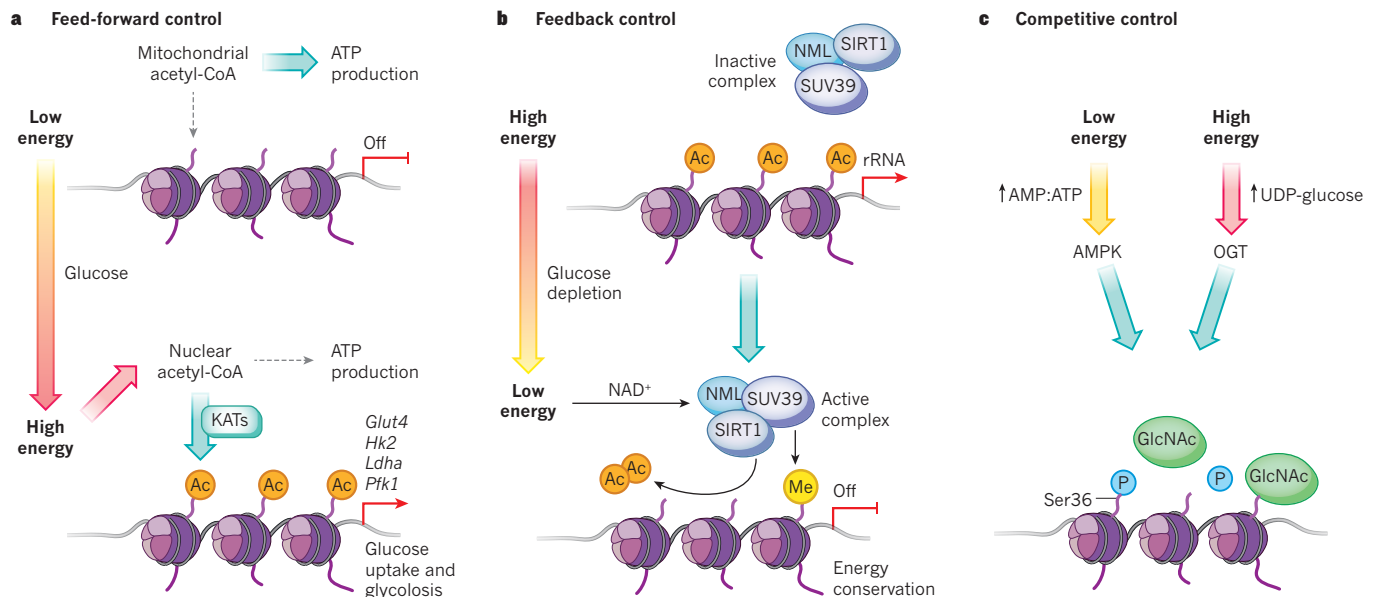


Figure 3 | Distinct modes of chromatin-mediated transcriptional control by intermediary metabolism products. **a**, Fasting leads to high acetyl-CoA levels in mitochondria and low levels in the nucleus. By contrast, during feeding or under high glucose conditions, acetyl-CoA from glycolysis is exported to the nuclear-cytosolic compartment. In the nucleus, increased acetyl-CoA activates KATs, which acetylate histones, thereby creating a permissive state for transcription of genes involved in glucose uptake (*Glut4*) and glycolysis (*Hk2*, *Ldha* and *Pfk1*). In this feed-forward loop acetyl-CoA transmits the signal of increased nutrient supply to a transcriptional program that accelerates the uptake and catabolism of glucose to acetyl-CoA as a building block for macromolecule synthesis. **b**, When cells transition from a high-energy environment to glucose depletion,

NAD⁺:NADH ratio rises. Activated SIRT1 deacetylates H3K9. In addition, the histone methyltransferase SUV39H1 transfers a methyl group to the liberated lysine residue, thereby locking chromatin in a silent state. This epigenetic switch preserves energy by suppressing the transcription of ribosomal RNA, a highly energy consuming process. **c**, A model for how two molecular sensors of intracellular energy levels might compete for unique histone marks of histone H2B. AMPK, which senses low energy levels, phosphorylates serine 36 (Ser36) of H2B, whereas OGT O-GlcNAcyates Ser36 of H2B when high glucose levels generate UDP-GlcNAc in the hexosamine biosynthetic pathway. Ac, acetylation; Me, methylation; P, phosphorylation; O-GlcNAc, O-linked N-acetylglucosamine; OGT, O-GlcNAc transferase; UDP, uridine diphosphate.

contrast to prokaryotes, multicellular organisms have evolved many additional regulatory steps of cellular growth, such as growth factor signalling and cell-cycle checkpoints, that may override changes induced by variations in acetyl-CoA concentrations¹¹. Nevertheless, these seminal discoveries in yeast and mammalian cells reveal that acetyl-CoA, KATs and HDACs may act as an integrative network, linking the elevated abundance of acetyl-CoA to downstream pathways of energy storage and proliferation.

Sirtuins as NAD⁺-sensing HDACs

As acetyl-CoA levels regulate histone acetylation during feeding conditions, not surprisingly, the reverse reaction, deacetylation, is also connected to intracellular energy levels. Among HDACs, sirtuins are the prime suspects as metabolic sensors. The beneficial effects of calorie

restriction on metabolic, locomotory and cognitive parameters are thought to depend on the sensing of the energy carrier NAD^+ by SIRT1 (refs 49, 50). However, questions remain as to whether NAD^+ levels fluctuate under a range of metabolic conditions, such as fasting and re-feeding or calorie restriction¹⁰. Various experimental models have demonstrated how changes in NAD^+ levels could influence the enzymatic activity of sirtuins, including fluctuations within microdomains at the chromatin level¹³. NAD^+ levels fluctuate substantially in a circadian manner, linking the peripheral clock to the transcriptional regulation of metabolism by epigenetic mechanisms involving SIRT1 (ref. 51). The core circadian clock machinery, BMAL1 and CLOCK, directly regulates expression of nicotinamide phosphoribosyltransferase (NAMPT), the rate-limiting enzyme of the NAD^+ salvage pathway in mice⁵². SIRT1 protein abundance is relatively stable, but its deacetylase activity depends on

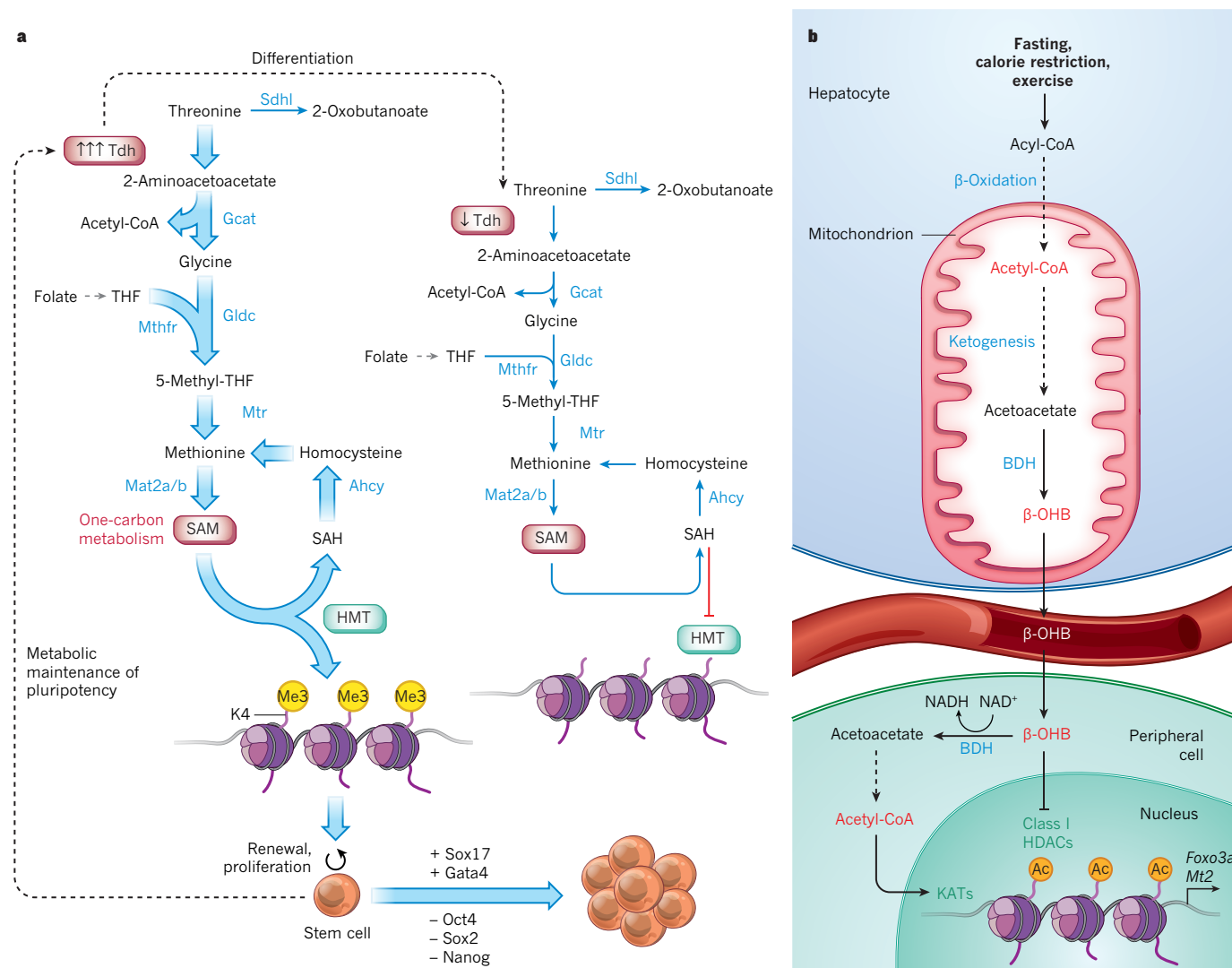


Figure 4 | Metabolite influencers of complex biological systems. a, Epigenetic control of pluripotency by threonine to S-adenosylmethionine (SAM) flux. Threonine dehydrogenase (Tdh) is about 200-fold enriched in embryonic stem cells compared with differentiated cells. Flux through this enzyme routes the amino acid threonine toward SAM production. After methyl transfer from SAM to a macromolecule, methionine is recovered in the one-carbon metabolism from S-adenosylhomocysteine (SAH) utilizing homocysteine and 5-methyl-tetrahydrofolate (THF). Deficiency of Tdh or depletion of threonine from the culture medium reduces the supply for SAM generation and leads to cellular differentiation as well as loss of stem cell markers. When the SAM:SAH ratio drops, SAH inhibits methyltransferases, leading to a reduction in methylation marks at lysine 4 of histone 3. **b,** Suppression of oxidative stress by the ketone

body β -hydroxybutyrate (β -OHB). Fasting, calorie restriction or exercise elevate production of the ketone body β -OHB in the liver. In peripheral cells, β -OHB can either directly inhibit histone deacetylases (HDACs) or increase nuclear acetyl-CoA levels. HDAC inhibition and stimulation of lysine acetyltransferases (KATs) by acetyl-CoA increases histone acetylation resulting in a permissive state for transcription of several genes (for example, *Foxo3a* and *Mt2*) of the oxidative damage response. Ahcy, S-adenosylhomocysteine hydrolase; BDH, β -hydroxybutyrate dehydrogenase; Gcat, glycine C-acetyltransferase; Glcd, glycine dehydrogenase (decarboxylating); HMT, histone methyltransferase; Mthfr, methylentetrahydrofolate reductase; Mtr, methyltetrahydrofolate-homocysteine methyltransferase; Mat2a/b, methionine adenosyltransferase 2a/b; Me3, trimethylated lysine residue; Sdh, serine dehydratase.

BOX 1

Transgenerational epigenetic inheritance of longevity

Organismal ageing is genetically controlled¹¹² and recent experiments suggest it is also regulated at the epigenetic level. Histone H3 methylation marks change during ageing in *D. melanogaster* and humans. In addition, genetic manipulations of components of demethylase or methyltransferase protein complexes affect life expectancy in *C. elegans* and *D. melanogaster*¹¹³. Remarkably, knockdown of the three COMPASS (complex proteins associated with Set1) members in *C. elegans* — ASH-2, WDR-5 and the histone methyltransferase SET-2 — not only extend lifespan in the worm in which the knockdown is induced but, unexpectedly, also in the next three generations¹¹⁴. Gene expression changes associated

with H3K4me3 complex deficiency are inherited at specific loci in the absence of a global reduction of H3K4me3 marks. Whether the incomplete chromatin reprogramming of a specific target gene or a larger transcriptional cluster is responsible for the lifespan extension has yet to be tested. Insulin/IGF1-signalling, TOR signalling and calorie restriction through sirtuins are mechanisms that regulate lifespan across species^{115,116}. Are environmental signals, including calorie restriction, able to induce stable epigenetic marks that influence complex traits such as the health span and lifespan of our children? The mechanisms behind such transgenerational effects are only just starting to be explored.

the presence of NAMPT to generate NAD⁺, and SIRT1 enzymatic activity oscillations correlate with the circadian production of NAD⁺. Notably, the circadian rhythmicity of NAMPT is abolished in *Clock* knockout mice. When a specific inhibitor, FK866, is used to block NAMPT activity, circadian oscillation of NAD⁺ levels are blunted and Sirt1 cyclic activity is lost⁵². The targets of SIRT1 include K9 and K14 of H3 at multiple loci of genes that oscillate under circadian control⁵³. Thus, NAD⁺ levels seem to directly regulate histone acetylation through the HDAC function of SIRT1 in a circadian manner, suggesting that NAD⁺ is indeed an important determinant for histone deacetylation.

SIRT1 regulates another homeostatic circuit that senses low energy levels. Ribosomal biogenesis is a highly energy-consuming process in eukaryotes, in particular in proliferative tissues, and its production rate is tightly linked to cellular energy levels to preserve cellular function under restricted substrate availability⁵⁴. A protein complex that contains nucleomethylin, the heterochromatin methyltransferase SUV39H1 and SIRT1 is a key regulator of this process. Glucose starvation leads to an increase in the NAD⁺:NADH ratio, SIRT1 activation and deacetylation of histone H3K9 at rDNA loci. This allows SUV39H1 to dimethylate H3K9. The concomitant structural chromatin switch from loosened euchromatin to condensed heterochromatin represses ribosomal biogenesis and protects cells from energy depletion (Fig. 3b).

SIRT6, which is also located in the nucleus, is linked to ageing by regulating telomere stability and inflammation through NF- κ B signalling. Deacetylation of histone H3K9 seems to be the chromatin modification that connects SIRT6 activity with these pathways of ageing^{55,56}. Loss of SIRT6 leads to progeria⁵⁷, whereas gain of function extends lifespan in male mice by 15%⁵⁸.

Whether fluctuation of NAD⁺ levels contributes to a change in activity of sirtuins remains controversial, particularly because classic studies show that intracellular NAD⁺ levels are kept constant even under changing metabolic conditions, including starvation^{10,59}. However, as discussed later, whole-cell or whole-tissue measurement of concentrations might mask compartment-specific differences of NAD⁺; for example, mitochondrial compared with nuclear. Importantly, recent experiments indicate that SIRT1 phosphorylated by adrenergic receptor signalling (through cAMP–PKA activation) becomes activated by lower NAD⁺ concentrations than unphosphorylated SIRT1 (ref. 60). In addition, nutritional supplementation with nicotinamide riboside, a precursor of NAD⁺, leads to elevated intracellular concentrations of NAD⁺, activation of Sirt1 and enhanced oxidative metabolism in mice⁶¹. A recent study identified a positive allosteric site adjacent to the enzyme domain of SIRT1 that is necessary for the enzyme to be activated by the red wine component resveratrol⁶². An intriguing possibility is that resveratrol mimics an endogenous allosteric metabolite for SIRT1 (ref. 63). Clearly, more work will be necessary to gain a full understanding of sirtuin regulation and the relative role of NAD⁺ fluctuations and other regulatory mechanisms.

Other connections

Adenosine-monophosphate-activated protein kinase (AMPK) is an atypical kinase that senses changes in AMP:ATP ratios and computes these into post-translational phosphorylation reactions of key regulators of whole-body and cellular energy levels, cellular stress response and cell-cycle control⁶⁴. Strikingly, AMPK phosphorylates histone H2B at serine 36 and thereby regulates the transcription of genes involved in the genotoxic and cellular metabolic stress response pathways⁶⁵ (Fig. 3c). Remarkably, the same residue of histone H2B can also be modified by O-GlcNAcylation by OGT, suggesting possible competition between the two modifications for H2BS36, although the biological relevance for such competition remains to be shown experimentally. Nevertheless, this could be a vivid example of a reciprocal regulation of a chromatin mark by two cellular energy sensors for low (AMPK) and high (OGT) energy levels.

Metabolic influencers of biological systems

Unique metabolites can also influence transcriptional programs, controlling stem-cell biology, cancer and ageing.

Methyl-donor requirement for pluripotency

The discovery by Yamanaka and colleagues that differentiated cells can be reprogrammed to induced pluripotent stem cells (iPSCs) has opened up the exciting prospect of tissue replacement therapies for degenerative diseases⁶⁶. Reprogramming involves the reversal of the epigenetic landscape acquired during development⁶⁷ (see the Review by Apostolou and Hochedlinger in this Insight⁶⁸). *In vitro* differentiation of mouse embryonic stem (ES) cells to embryonic bodies is associated with marked compositional changes in their metabolome, including an accumulation of metabolites related to one-carbon metabolism, threonine metabolism and acetyl-CoA generation⁶⁹. Withdrawal of threonine or genetic deletion of threonine dehydrogenase (Tdh), which converts threonine to 2-amino-3-ketobutyrate as a supply branch for one-carbon metabolism, completely represses proliferation of mouse ES cells and commits them to differentiation. This effect is unique to threonine and does not occur when any other amino acid is experimentally omitted. The distinct metabolic composition of pluripotent cells is achieved by a marked upregulation of Tdh in mouse ES cells (200-fold more than mouse embryonic fibroblasts (MEFs))⁶⁹. A recent study links these findings to levels of H3K4me3 (ref. 70) (Fig. 4a). When MEFs are reprogrammed to iPSCs, they acquire a metabolite pattern that resembles mouse ES cells, including a pathway enrichment of enzymatic reactions that feed one-carbon metabolism. This pathway starts with the amino acid threonine, which is converted in two steps to glycine and acetyl-CoA. Glycine donates a methyl group to derive 5-methyl-tetrahydrofolate from tetrahydrofolate, leading to increased levels of SAM. An unidentified methyltransferase responds to the increased flux in the threonine-SAM pathway and specifically trimethylates H3K4. Notably, the product of the methylation reaction, SAH, acts as a negative feedback regulator for methyltransferases, further

supporting the model that SAM:SAH ratio regulates histone methylation. Like withdrawal of threonine, inhibition of Tdh in mouse ES cells results in decreased SAM:SAH ratios and a selective reduction of H3K4me2 and H3K4me3. The biological consequence are an impaired growth of ES cells and loss of pluripotency, suggesting that rapid metabolic flux from threonine to SAM is required to maintain a chromatin signature associated with pluripotency. The relevance of this regulatory system in human stem cells is currently unclear.

Oncometabolites and a metabolite tumour suppressor

Epigenetic processes are involved in the causation and progression of many cancers. Systematic profiling of the 'cancer epigenome' over the past decade has determined a large number of altered chromatin marks across the genome that are associated with tumours that potentially regulate gene expression and genome stability⁷¹. Increasingly, this rapidly evolving field is merging with cancer metabolism. Nearly all proliferating cancer cells adopt a metabolic state tailored to their specific energetic needs for rapid proliferation and growth, a metabolic condition termed the Warburg effect⁷². In brief, cancer cells shift from a mitochondrial oxidative metabolism to an aerobic glycolytic metabolism that provides small metabolites, including acetyl-CoA, as building blocks for the increased macromolecule synthesis that is necessary for rapid division and growth. For example, the oncogenic pyruvate kinase isoform M2 is a glycolytic enzyme that promotes aerobic glycolysis in cancer cells⁷³.

Emerging evidence indicates that metabolites can alter tumour properties through epigenetic mechanisms. In numerous cancers, predominantly glioblastoma and acute myeloid leukaemia, isocitrate dehydrogenase isoforms 1 and 2 (IDH1 and IDH2) contain hot spots for somatic mutations in or near their enzymatic pockets⁷⁴. Intriguingly, only one of the two alleles is affected by the mutation, giving rise to a heterogeneous pool of wild-type and mutant proteins. Interestingly, the mutated form gains a neomorphic enzymatic property and converts the normal product of IDH, α -KG, in a sequential reaction to 2-hydroxyglutarate, resulting in drastically elevated tissue concentrations³⁹. 2-Hydroxyglutarate is a competitive inhibitor of several α -KG-dependent dioxygenases, including the histone demethylase KDM4C and the TET hydroxylases. Thus, the inhibition of such dioxygenases by 2-hydroxyglutarate might help to maintain cells in an undifferentiated state, potentially priming them for malignant transformation^{75,76}.

Butyrate is a potent inhibitor of several HDACs (with a half-maximal inhibitory concentration (IC_{50}) of 90 μ M in HT-29 colon-carcinoma-derived cells)⁷⁷. This short-chain fatty acid is produced at high concentrations in the lumen of the colon from dietary fibres and provides the main energy source for colonocytes, in which it is rapidly oxidized to acetyl-CoA^{31,78}. A diet rich in fibre is thought to prevent colitis and colon cancer in humans⁷⁹. Recent findings support the model that butyrate regulates transcriptional programs involved in proliferation of colon cells⁸⁰. High levels of butyrate (about 5 mM), which occur at the epithelial wall of the proximal colon, inhibit the enzymatic activity of class I/II HDACs. Loss of HDAC activity increases lysine acetylation of H3 and regulates a transcriptional program that decreases proliferation. By contrast, due to diffusion dynamics, cells in the colon crypts are exposed to 10-fold lower concentrations of butyrate (around 0.5 μ M). At this low concentration, negligible effects on HDAC activity occur *in vitro*. Nevertheless, radiotracer flux studies show that butyrate-derived acetyl-CoA is removed from the mitochondria by the citrate shuttle and contributes to histone acetylation in a reaction that is dependent on ATP-citrate lyase. This ATP-citrate-lyase-dependent histone acetylation targets a different gene subset than high-butyrate, HDAC-dependent, histone acetylation and includes key regulators that promote cellular proliferation. The proposed purpose of this concentration-dependent transcriptional control is to stimulate the growth of progenitors in the colon crypts, while maintaining quiescence of luminal cells⁸⁰. The metabolism of colon cancer cells shifts to aerobic glycolysis and throttle β -oxidation, leading to increased butyrate levels that match the high luminal concentrations. The resulting direct

inhibition of HDACs activates anti-proliferative and apoptotic genes and may explain the anti-cancer effects that have been described for butyrate⁸¹. Whether a fibre-rich diet sufficiently increases butyrate levels beyond the normal levels of these short-chain fatty acids to explain an anti-cancer effect by HDAC inhibition remains to be addressed.

Ketone bodies, oxidative stress and ageing

Ketone bodies are produced mainly in the liver as alternative energy substrates when the glucose supply drops critically during fasting⁸². Neurons and other peripheral tissues avidly resorb and consume ketone bodies as a carbon source for ATP production. Prolonged fasting, calorie restriction, strenuous exercise or ketogenic diets are associated with increases in serum concentrations of the ketone body β -hydroxybutyrate (β -OHB) from micromolar to millimolar concentrations (about 2–6 mM). A protective effect of ketogenic diets on survival of neurons in models of Alzheimer's and Parkinson's disease was described 10 years ago⁸³. In addition, a reduction of reactive oxygen species, by-products of mitochondrial oxidative metabolism, is observed during calorie restriction and ketogenic diets⁸⁴. β -OHB is structurally similar to butyrate. Most intriguingly, β -OHB also acts as an inhibitor of HDACs⁸⁵ (Fig. 4b). Inhibition of HDAC1, 3 and 4 by this endogenous metabolite increases acetylation of histone H3K9 and K14 and establishes a permissive chromatin configuration for expression of several key components of the oxidative damage response in the kidney. These include the longevity-associated transcription factor Foxo3a and the metallothionein Mt2. This example illustrates how β -OHB, a metabolite used as a circulating glucose-sparing energy source, can also serve as a signalling molecule that regulates a unique transcriptional program associated with a specific metabolic condition (fasting or exercise).

Persistent epigenetic changes induced by nutrients

The influence of DNA and chromatin modifications on transcription regulation may help to translate transient metabolic states into more stable transcriptional states that persist and affect phenotypes over extended periods of time. An example of this hypothesis has come from studies in the honeybee (*Apis mellifera*). Feeding 'royal jelly' to the future queen has a fundamental impact on adult morphology, behaviour and longevity⁸⁶ and also changes DNA methylation and gene expression patterns. Genetically identical larvae that do not receive the royal jelly develop into sterile workers. How royal-jelly feeding leads to the changes in these complex traits and gene expression patterns is unknown.

During vertebrate development, DNA and histone methylation patterns become stabilized and are retained throughout life. Genes that are subject to changes in initially fixed epigenetic states are termed metastable epialleles. This epigenetic drift occurs in response to intrinsic and environmental factors and has been described in monozygotic twins^{87,88}. Although the epigenome is indistinguishable during early life, the overall content and genomic distribution of both DNA methylation and histone acetylation differ substantially among monozygotic twins⁸⁷. Relatively little is known about how chromatin modifications acquired in post-mitotic, adult tissues become stabilized and influence transcription and disease risks for extended periods of time. Diabetic patients who maintain intense glucose control and near-to-normal glycosylated HbA1c, remain at increased risk of macrovascular complications and diabetic organ damage even years after the initial diagnosis^{89,90}. The mechanisms of this 'glycaemic memory' are not well understood, but persistence of a chronic inflammation is believed to have a role⁹¹. Primary human aortic cells show changes to H3K9 and H3K14, and to H3K4me2 and H3K4me3 in response to increased glucose exposure⁹². Interestingly, when endothelial cells are isolated from diabetic mice, or exposed to high-glucose bouts before switching to a low-glucose medium, long-term activation of several key inflammatory genes persists. For example, chronic pro-inflammatory gene expression correlates with an increase of H3K4 methylation and the suppression of H3K9me2 and me3 marks at the NF- κ B-p65 promoter^{93,94}. Given that cardiovascular disease remains the number one cause of death in industrialized countries and the rising epidemic of obesity and diabetes,

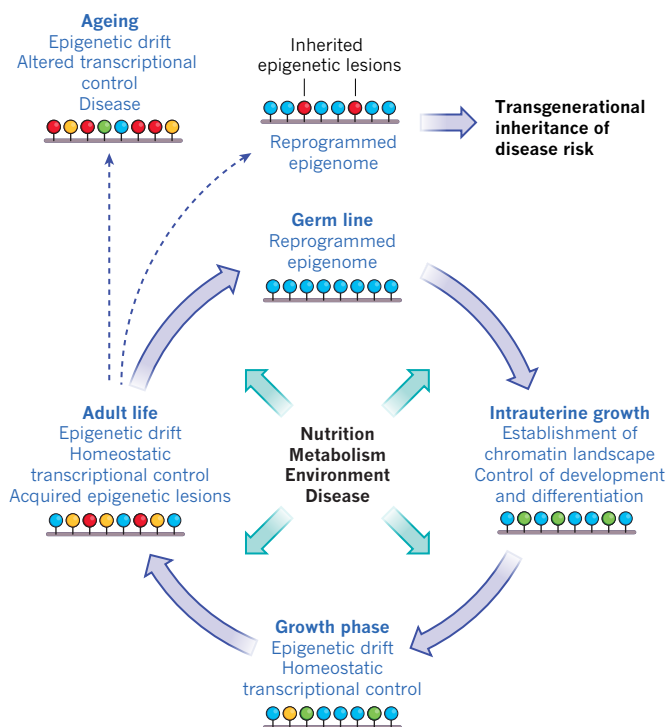


Figure 5 | Epigenetic drift and transgenerational inheritance of disease risks. DNA methylation and histone marks are established during embryonic development to maintain cell lineage commitment. After birth the chromatin landscape retains a dynamic configuration throughout life. Changes of chromatin marks within a gene locus, termed epigenetic drift, occur in response to nutritional, metabolic, environmental or pathological signals and are part of homeostatic adaptations. When adverse epigenetic drift compromises a cell's capacity to adequately respond to challenges, disease susceptibility increases stochastically. Under some circumstances epigenetic marks may escape epigenetic reprogramming during gametogenesis and be inherited by subsequent generations. Transgenerational inheritance of epigenetic regulation would contribute to disease susceptibility by transmitting an acquired epigenetic predisposition to the next generation independently of genetically inherited risk factors.

the implications of a glycaemic memory seem relatively underappreciated. *In vivo* models that can be used to capture epigenetic changes in response to temporarily restricted exposure to 'epigenetically toxic' metabolites such as glucose are urgently needed.

Inheritance of epigenetic traits

Changes in the chromatin landscape that accumulate during development and adulthood are actively reprogrammed during gametogenesis¹⁶. Despite this apparent ‘new start for a new life’, some epigenetic information in response to environmental factors resists reprogramming and is thereby transmitted to the next generation⁸⁸.

The first example of transgenerational inheritance of longevity was recently reported in *C. elegans*. This multifactorial trait, increased lifespan, is associated with the loss of specific histone methylation marks and persists for three generations (Box 1). The study of epigenetic inheritance of complex traits, such as diabetes, obesity and ageing, is a rapidly evolving, exciting new frontier; however, evidence for an inheritable component of epigenetic gene regulation in response to nutrients is mostly correlative and little mechanistic insight exists so far. Two classic epidemiological studies analysed how exposure to undernutrition or overnutrition during intrauterine development relates to the risk of metabolic disease in first- and second-generation descendants^{32,95}. The first study found that children of mothers who endured the severe Dutch famine of 1944 had a low birth weight, but significantly increased risk of adiposity in later life⁹⁶. The second cohort study investigated the risk for metabolic disease in response to food supply fluctuations in the remote region of Överkalix

in northern Sweden. Notably, the prevalence of cardiovascular disease and diabetes was elevated, and longevity decreased, in grandsons of grandfathers who were well fed between 8 and 12 years of age compared with those for whom availability to food was poor between the same ages⁹⁷. In mammals, several experimental studies have established a generational transmission of adverse metabolic phenotypes in response to undernutrition or overnutrition in the parent generation⁹⁵. These results have to be regarded with caution, however, because no mechanistic link between the pathological outcomes and specific chromatin-based mechanisms have been shown.

A complication for studying epigenetic transmission in mammals is to separate direct effects on the maternal germ line from intra-uterine exposure of the offspring. Several studies have tested the contribution of paternal inheritance in response to changes in metabolic conditions. Pre-mating exposure of male mice to fasting affects serum glucose in the offspring of both genders⁹⁸. Both male and female offspring show an upregulation of genes involved in lipid and cholesterol biosynthesis in the liver in response to a paternal low-protein diet⁹⁹. Modest but reproducible changes in DNA methylation in an enhancer region of *Ppara* — which encodes a nuclear receptor that functions as a key regulator of adaptive fasting response in mice, including gluconeogenic and ketogenic gene expression — have been described¹⁰⁰. A landmark study in rats links paternal diet to insulin-producing β -cell dysfunction specifically in female offspring¹⁰¹. Chronic feeding with a high-fat diet alters expression of islet genes associated with glucose metabolism and impairs glucose tolerance as well as insulin sensitivity in the next generation¹⁰¹. Again, these interesting findings are so far correlative and mechanistic insights are required to establish a clear link between nutritional conditions and heritable chromatin-mediated changes in gene regulation.

Perspective and challenges

The discovery of endogenous metabolites that help to drive transcriptional programs as rheostat inputs or feed-forward signals for cell-fate decisions is a fascinating example of the integration of multiple cellular functions. All of the connections that exist between metabolites and transcriptional regulators have probably not been discovered yet. The generation of large data sets that provide information on DNA methylation, histone modification and nucleosome positioning on a genome-wide basis, particularly the ENCODE project¹⁰², will be important tools in this effort. In addition, metabolomics, proteomics and whole-genome sequencing methods allow the measurement of compartment-specific metabolic states, which can then be correlated with chromatin states and gene expression.

However, given the complex traits that are regulated by chromatin-mediated mechanisms, it will be difficult to extrapolate discoveries from simple experimental models such as yeast and mammalian cell culture systems to the complexity of a whole organism. Future efforts will therefore need to focus on the role of organ-, age- and disease-specific metabolic changes in the regulation of specific enzymes.

Metabolic epigenetic reprogramming

These findings also raise a number of crucial questions. For example, how do specific dietary interventions, such as calorie restriction or ketogenic diets, affect chromatin and transcription and lead to beneficial effects on metabolic health? In addition, do changes in diet or changes in metabolism that are associated with obesity lead to epigenetic drift with the potential of transgenerational inheritance? The growing epidemic of obesity and metabolic disease in the Western world is of particular concern. If metabolic changes associated with obesity indeed lead to epigenetic changes, some of those could be inherited transgenerationally and lead to epigenetic predisposition to metabolic disease in subsequent generations. If this is the case, a potential vicious cycle develops (Fig. 5). In this model of ‘inherited genetic drift’, epigenetic modifications acquired during chronological ageing reduce the capacity for homeostatic responses¹⁰³. Some of these marks may be passed on to the next generation (or generations) through incomplete reprogramming in the germ line⁸⁸.

The high prevalence of obesity and cardiovascular disease is thought to be responsible for the recently observed decrease in life expectancy¹⁰⁴. Understanding the influence of nutrients, metabolites and other environmental factors on the metabolic landscape and its influence on the epigenome will probably open up new therapeutic strategies. Could the DNA methylation and histone modification patterns be reprogrammed before conception or during development to erase the altered epigenome? The enzymes controlling epigenetic modifications will probably represent a fertile ground for the discovery of new drug targets, and clinical trials of histone modifier drugs are under way for various diseases¹⁰⁵ (see Review by Helin and Dhanak in this Insight¹⁰⁶).

Biomarker sensors

As already discussed for NAD⁺ and SIRT1, a full understanding of the nexus between intermediary metabolites and chromatin regulators will require the development of highly sensitive and selective sensors that measure metabolite concentrations in different organs and cellular compartments. New tools have been developed to measure NADPH¹⁰⁷, SAM¹⁰⁸, malonyl-CoA¹⁰⁹ and activated AMPK¹¹⁰. The use of these fluorescence-based reporters will allow key questions to be answered. For example, comparing the nuclear, cytosolic and mitochondrial concentrations of distinct metabolites or measurement of tissue gradients of metabolites. A peroxidase sensor, HyPer, has shown that leukocytes are recruited to a wound zone by a local hydrogen peroxide gradient *in vivo*¹¹¹. Similarly, metabolite sensors could be used in translucent zebrafish (*Danio rerio*) to measure the distribution of epigenetic regulator metabolites in various disease models across tissue gradients or with subcellular resolution.

The previously predicted regulation of epigenetic programs by metabolites is emerging as an important mechanism of biological integration of distinct cellular functions. Much remains to be discovered, and the study of crosstalk between metabolism and epigenetic regulators will probably bring a more integrated understanding of cellular and organismal functioning and, possibly, new therapeutic opportunities. ■

Received 15 March; accepted 16 August 2013.

- Feinberg, A. P. Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440 (2007).
- Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
- Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
- This landmark review proposes for the first time the existence of a complex histone code that encodes information as the underlying logic of epigenetic transcriptional control.**
- Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
- Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
- Kornberg, R. D. Structure of chromatin. *Annu. Rev. Biochem.* **46**, 931–954 (1977).
- Bell, O., Tiwari, V. K., Thomä, N. H. & Schübeler, D. Determinants and dynamics of genome accessibility. *Nature Rev. Genet.* **12**, 554–564 (2011).
- Musselman, C. A., Lalonde, M.-E., Cote, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nature Struct. Mol. Biol.* **19**, 1218–1227 (2012).
- Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
- Kaelin, W. G. & McKnight, S. L. Influence of metabolism on epigenetics and disease. *Cell* **153**, 56–69 (2013).
- Lu, C. & Thompson, C. B. Metabolic regulation of epigenetics. *Cell Metab.* **16**, 9–17 (2012).
- Wellen, K. E. & Thompson, C. B. A two-way street: reciprocal regulation of metabolism and signalling. *Nature Rev. Mol. Cell Biol.* **13**, 270–276 (2012).
- Katada, S., Imhof, A. & Sassone-Corsi, P. Connecting threads: epigenetics and metabolism. *Cell* **148**, 24–28 (2012).
- Teperino, R., Schoonjans, K. & Auwerx, J. Histone methyl transferases and demethylases; can they link metabolism and transcription? *Cell Metab.* **12**, 321–327 (2010).
- Bird, A. Perceptions of epigenetics. *Nature* **447**, 396–398 (2007).
- Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature Rev. Genet.* **9**, 129–140 (2008).
- Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).
- Hirschey, M. D. *et al.* SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature* **464**, 121–125 (2010).
- Newman, J. C., He, W. & Verdin, E. Mitochondrial protein acylation and intermediary metabolism: regulation by sirtuins and implications for metabolic disease. *J. Biol. Chem.* **287**, 42436–42443 (2012).
- He, W., Newman, J. C., Wang, M. Z., Ho, L. & Verdin, E. Mitochondrial sirtuins: regulators of protein acylation and metabolism. *Trends Endocrinol. Metab.* **23**, 467–476 (2012).
- Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Forneris, F., Binda, C., Vanoni, M. A., Mattevi, A. & Battaglioli, E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett.* **579**, 2203–2207 (2005).
- Loenarz, C. & Schofield, C. J. Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chem. Biol.* **4**, 152–156 (2008).
- Sakabe, K., Wang, Z. & Hart, G. W. β -N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl Acad. Sci. USA* **107**, 19915–19920 (2010).
- Hart, G. W. G., Housley, M. P. M. & Slawson, C. C. Cycling of O-linked β -N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**, 1017–1022 (2007).
- Hanover, J. A. J., Krause, M. W. M. & Love, D. C. D. Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation. *Nature Rev. Mol. Cell Biol.* **13**, 312–321 (2012).
- Fong, J. J. *et al.* β -N-Acetylglucosamine (O-GlcNAc) is a novel regulator of mitosis-specific phosphorylations on histone H3. *J. Biol. Chem.* **287**, 12195–12203 (2012).
- Rodgers, J. T. *et al.* Nutrient control of glucose homeostasis through a complex of PGC-1 α and SIRT1. *Nature* **434**, 113–118 (2005).
- Bird, A. P. Functions for DNA methylation in vertebrates. *Cold Spring Harb. Symp. Quant. Biol.* **58**, 281–285 (1993).
- Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013).
- Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nature Biotechnol.* **28**, 1057–1068 (2010).
- Kirchner, H., Osler, M. E., Krook, A. & Zierath, J. R. Epigenetic flexibility in metabolic regulation: disease cause and prevention? *Trends Cell Biol.* **23**, 203–209 (2013).
- Barrès, R. *et al.* Non-CpG methylation of the PGC-1 α promoter through DNMT3B controls mitochondrial density. *Cell Metab.* **10**, 189–198 (2009).
- Barrès, R. *et al.* Weight loss after gastric bypass surgery in human obesity remodels promoter methylation. *Cell Rep.* **3**, 1020–1027 (2013).
- Anderson, O. S., Sant, K. E. & Dolinoy, D. C. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J. Nutr. Biochem.* **23**, 853–859 (2012).
- Xiao, M. *et al.* Inhibition of α -KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* **26**, 1326–1338 (2012).
- Ladurner, A. G. Rheostat control of gene expression by metabolites. *Mol. Cell* **24**, 1–11 (2006).
- Shi, Y. & Shi, Y. Metabolic enzymes and coenzymes in transcription — a direct link between metabolism and transcription? *Trends Genet.* **20**, 445–452 (2004).
- Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).
- This breakthrough article describes the neomorphic enzyme activity of cancer-associated mutant IDH1 leading to tissue accumulation of the oncometabolite 2-hydroxyglutarate.**
- Albaugh, B. N., Arnold, K. M. & Denu, J. M. KAT(ching) metabolism by the tail: insight into the links between lysine acetyltransferases and metabolism. *ChemBioChem* **12**, 290–298 (2011).
- Rardin, M. J. *et al.* Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways. *Proc. Natl Acad. Sci. USA* **110**, 6601–6606 (2013).
- Hebert, A. S. *et al.* Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. *Mol. Cell* **49**, 186–199 (2012).
- Tu, B. P. & McKnight, S. L. Metabolic cycles as an underlying basis of biological oscillations. *Nature Rev. Mol. Cell Biol.* **7**, 696–701 (2006).
- Tu, B. P. *et al.* Cyclic changes in metabolic state during the life of a yeast cell. *Proc. Natl Acad. Sci. USA* **104**, 16886–16891 (2007).
- Berndsen, C. E. & Denu, J. M. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* **18**, 682–689 (2008).
- Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Mol. Cell* **42**, 426–437 (2011).
- Takahashi, H., McCaffery, J. M., Irizarry, R. A. & Boeke, J. D. Nucleocytoplasmic acetyl-coenzyme a synthetase is required for histone acetylation and global transcription. *Mol. Cell* **23**, 207–217 (2006).
- Wellen, K. E. *et al.* ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076–1080 (2009).
- References 46 to 48 show that acetyl-CoA increases histone acetylation in yeast and mammalian cells when its abundance is enriched in the nucleocytoplasmic compartment of a cell.**
- Haigis, M. C. & Guarente, L. P. Mammalian sirtuins — emerging roles in physiology, aging, and calorie restriction. *Genes Dev.* **20**, 2913–2921 (2006).
- Houtkooper, R. H., Pirinen, E. & Auwerx, J. Sirtuins as regulators of metabolism and healthspan. *Nature Rev. Mol. Cell Biol.* **13**, 225–238 (2012).
- Sassone-Corsi, P. NAD⁺, a circadian metabolite with an epigenetic twist. *Endocrinology* **153**, 1–5 (2012).

52. Nakahata, Y., Sahar, S., Astarita, G., Kaluzova, M. & Sassone-Corsi, P. Circadian control of the NAD⁺ salvage pathway by CLOCK–SIRT1. *Science* **324**, 654–657 (2009).
53. Nakahata, Y. *et al.* The NAD⁺-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. *Cell* **134**, 329–340 (2008).
54. Murayama, A. *et al.* Epigenetic control of rDNA loci in response to intracellular energy status. *Cell* **133**, 627–639 (2008).
55. Michishita, E. *et al.* SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* **452**, 492–496 (2008).
56. Kawahara, T. L. A. *et al.* SIRT6 links histone H3 lysine 9 deacetylation to NF- κ B-dependent gene expression and organismal life span. *Cell* **136**, 62–74 (2009).
57. Mostoslavsky, R. *et al.* Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* **124**, 315–329 (2006).
58. Kanfi, Y. *et al.* The sirtuin SIRT6 regulates lifespan in male mice. *Nature* **483**, 218–221 (2012).
59. Veech, R. L. R., Eggleston, L. V. L. & Krebs, H. A. H. The redox state of free nicotinamide-adenine dinucleotide phosphate in the cytoplasm of rat liver. *Biochem. J.* **115**, 609–619 (1969).
60. Gerhart-Hines, Z. *et al.* The cAMP/PKA pathway rapidly activates SIRT1 to promote fatty acid oxidation independently of changes in NAD. *Mol. Cell* **44**, 851–863 (2011).
61. Cantó, C. *et al.* The NAD⁺ precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. *Cell Metab.* **15**, 838–847 (2012).
62. Hubbard, B. P. *et al.* Evidence for a common mechanism of SIRT1 regulation by allosteric activators. *Science* **339**, 1216–1219 (2013).
63. Gut, P. & Verdin, E. Rejuvenating SIRT1 activators. *Cell Metab.* **17**, 635–637 (2013).
64. Hardie, D. G. D., Ross, F. A. F. & Hawley, S. A. S. AMPK: a nutrient and energy sensor that maintains energy homeostasis. *Nature Rev. Mol. Cell Biol.* **13**, 251–262 (2012).
65. Bungard, D. *et al.* Signaling kinase AMPK activates stress-promoted transcription via histone H2B phosphorylation. *Science* **329**, 1201–1205 (2010).
66. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
67. Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**, 508–525 (2010).
68. Apostolou, E. & Hochedlinger, K. Chromatin dynamics during cellular reprogramming. *Nature* **502**, 462–471 (2013).
69. Wang, J. *et al.* Dependence of mouse embryonic stem cells on threonine catabolism. *Science* **325**, 435–439 (2009).
70. Shyh-Chang, N. *et al.* Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science* **339**, 222–226 (2013). **References 69 and 70 describe the dependence of pluripotency on the generation of S-adenosylmethionine by threonine catabolism in embryonic stem cells.**
71. Baylín, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nature Rev. Cancer* **11**, 726–734 (2011).
72. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
73. Vander Heiden, M. G. *et al.* Evidence for an alternative glycolytic pathway in rapidly proliferating cells. *Science* **329**, 1492–1499 (2010).
74. Sturm, D. *et al.* Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* **22**, 425–437 (2012).
75. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479–483 (2012).
76. Dang, L., Jin, S. & Su, S. M. IDH mutations in glioma and acute myeloid leukemia. *Trends Mol. Med.* **16**, 387–397 (2010).
77. Waldecker, M., Kautenburger, T., Daumann, H., Busch, C. & Schrenk, D. Inhibition of histone-deacetylase activity by short-chain fatty acids and some polyphenol metabolites formed in the colon. *J. Nutr. Biochem.* **19**, 587–593 (2008).
78. Roediger, W. E. Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* **83**, 424–429 (1982).
79. Kim, Y. S. & Milner, J. A. Dietary modulation of colon cancer risk. *J. Nutr.* **137**, 2576S–2579S (2007).
80. Donohoe, D. R. *et al.* The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol. Cell* **48**, 612–626 (2012).
81. Comalada, M. *et al.* The effects of short-chain fatty acids on colon epithelial proliferation and survival depend on the cellular phenotype. *J. Cancer Res. Clin. Oncol.* **132**, 487–497 (2006).
82. Cahill, G. F. Fuel metabolism in starvation. *Annu. Rev. Nutr.* **26**, 1–22 (2006).
83. Kashiwaya, Y. *et al.* D- β -hydroxybutyrate protects neurons in models of Alzheimer's and Parkinson's disease. *Proc. Natl Acad. Sci. USA* **97**, 5440–5444 (2000).
84. Maalouf, M., Rho, J. M. & Mattson, M. P. The neuroprotective properties of calorie restriction, the ketogenic diet, and ketone bodies. *Brain Res. Rev.* **59**, 293–315 (2009).
85. Shimazu, T. *et al.* Suppression of oxidative stress by β -hydroxybutyrate, an endogenous histone deacetylase inhibitor. *Science* **339**, 211–214 (2013). **This study identifies the ketone body β -hydroxybutyrate as an endogenous inhibitor of HDACs and protective agent against oxidative damage.**
86. Kucharski, R., Maleszka, J., Foret, S. & Maleszka, R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**, 1827–1830 (2008).
87. Fraga, M. F. M. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci. USA* **102**, 10604–10609 (2005). **This study in humans shows that epigenetic marks increasingly differ during adult life of monozygotic twins.**
88. Feil, R. & Fraga, M. F. Epigenetics and the environment: emerging patterns and implications. *Nature Rev. Genet.* **13**, 97–109 (2011).
89. Nathan, D. M. *et al.* Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N. Engl. J. Med.* **353**, 2643–2653 (2005).
90. Patel, A. *et al.* Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N. Engl. J. Med.* **358**, 2560–2572 (2008).
91. El-Osta, A. Glycemic memory. *Curr. Opin. Lipidol.* **23**, 24–29 (2012).
92. Piroola, L. *et al.* Genome-wide analysis distinguishes hyperglycemia regulated epigenetic signatures of primary vascular cells. *Genome Res.* **21**, 1601–1615 (2011).
93. Villeneuve, L. M. *et al.* Epigenetic histone H3 lysine 9 methylation in metabolic memory and inflammatory phenotype of vascular smooth muscle cells in diabetes. *Proc. Natl Acad. Sci. USA* **105**, 9047–9052 (2008).
94. Brasacchio, D. *et al.* Hyperglycemia induces a dynamic cooperativity of histone methylase and demethylase enzymes associated with gene-activating epigenetic marks that coexist on the lysine tail. *Diabetes* **58**, 1229–1236 (2009).
95. Patti, M.-E. Intergenerational programming of metabolic disease: evidence from human populations and experimental animal models. *Cell. Mol. Life Sci.* **70**, 1597–1608 (2013).
96. Stein, A. D. *et al.* Anthropometric measures in middle age after exposure to famine during gestation: evidence from the Dutch famine. *Am. J. Clin. Nutr.* **85**, 869–876 (2007).
97. Kaati, G., Bygren, L. O., Pembrey, M. & Sjöström, M. Transgenerational response to nutrition, early life circumstances and longevity. *Eur. J. Hum. Genet.* **15**, 784–790 (2007). **References 96 and 97 are classic epidemiological studies that show a relationship between nutritional status of mothers and cardiovascular and metabolic disease risk in the following generations.**
98. Anderson, L. M. *et al.* Preconceptional fasting of fathers alters serum glucose in offspring of mice. *Nutrition* **22**, 327–331 (2006).
99. Carone, B. R. *et al.* Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* **143**, 1084–1096 (2010).
100. Kersten, S. *et al.* Peroxisome proliferator-activated receptor α mediates the adaptive response to fasting. *J. Clin. Invest.* **103**, 1489–1498 (1999).
101. Ng, S.-F. *et al.* Chronic high-fat diet in fathers programs β -cell dysfunction in female rat offspring. *Nature* **467**, 963–966 (2010).
102. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
103. Rando, T. A. & Chang, H. Y. Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. *Cell* **148**, 46–57 (2012).
104. Olshansky, S. J. *et al.* A potential decline in life expectancy in the United States in the 21st century. *N. Engl. J. Med.* **352**, 1138–1145 (2005).
105. Arrowsmith, C. H., Bountra, C., Fish, P. V., Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nature Rev. Drug Discov.* **11**, 384–400 (2012).
106. Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488 (2013).
107. Zhang, Q., Piston, D. W. & Goodman, R. H. Regulation of corepressor function by nuclear NADH. *Science* **295**, 1895–1897 (2002).
108. Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).
109. Ellis, J. M. & Wolfgang, M. J. A genetically encoded metabolite sensor for malonyl-CoA. *Chem. Biol.* **19**, 1333–1339 (2012).
110. Tsou, P., Bin, B., Zheng, H., Hsu, C.-H., Sasaki, A. T. & Cantley, L. C. A fluorescent reporter of AMPK activity and cellular energy stress. *Cell Metab.* **13**, 476–486 (2011).
111. Niethammer, P., Grabher, C., Look, A. T. & Mitchison, T. J. A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish. *Nature* **459**, 996–999 (2009).
112. Kenyon, C. J. The genetics of ageing. *Nature* **464**, 504–512 (2010).
113. Han, S. & Brunet, A. Histone methylation makes its mark on longevity. *Trends Cell Biol.* **22**, 42–49 (2012).
114. Greer, E. L. *et al.* Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* **479**, 365–371 (2011).
115. Fontana, L., Partridge, L. & Longo, V. D. Extending healthy life span — from yeast to humans. *Science* **328**, 321–326 (2010).
116. Bordonie, L. & Guarente, L. Calorie restriction, SIRT1 and metabolism: understanding longevity. *Nature Rev. Mol. Cell Biol.* **6**, 298–305 (2005).

Acknowledgements We thank J. Carroll for graphics, G. Howard and A.-L. Lucido for editorial assistance, V. Fonseca and P. Cruz for administrative assistance. The authors are supported by funds from the Gladstone Institutes.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of this paper at go.nature.com/tagbox. Readers are welcome to comment on the online version of this article at go.nature.com/tagbox. Correspondence should be addressed to E.V. (everdin@gladstone.ucsf.edu).

Topology of mammalian developmental enhancers and their regulatory landscapes

Wouter de Laat¹ & Denis Duboule^{2,3}

How a complex animal can arise from a fertilized egg is one of the oldest and most fascinating questions of biology, the answer to which is encoded in the genome. Body shape and organ development, and their integration into a functional organism all depend on the precise expression of genes in space and time. The orchestration of transcription relies mostly on surrounding control sequences such as enhancers, millions of which form complex regulatory landscapes in the non-coding genome. Recent research shows that high-order chromosome structures make an important contribution to enhancer functionality by triggering their physical interactions with target genes.

Access to animal genome sequences has revealed that the level of complexity of an organism does not relate to its number of genes. Mammals are more complex in morphology and behaviour than roundworms, but their genomes both contain around 20,000 genes. Various parameters can contribute to increased complexity, such as the extent of protein modifications or the diversity of splicing patterns. Pleiotropy is another possible contributor, whereby genes acquire multiple functional tasks at different times and places either during development or in adult life. In this case, gene regulation, rather than function, had to evolve to associate regulatory alternatives to particular genes. Although gene transcription is initiated at promoters, which recruit the basal transcription machinery, these sequences have little impact on transcription control during development and hence this latter task mostly relies on enhancers¹.

Enhancers are sequence modules that contain binding motifs for transcription factors. They are preferentially located in the non-coding part of the genome, at various distances from their target genes^{2,3}. In mammals, more than 95% of the genome is non-coding and large gene deserts can sometimes span several megabases. The recent development of high-throughput methods has made it possible to systematically search for enhancers; millions of such regulatory modules have been predicted⁴, with 40% of our genome now estimated to carry some regulatory potential⁵. The importance of enhancers for normal development and disease is further underscored by the fact that disease-associated single nucleotide polymorphisms (SNPs) often co-localize with these modules⁶. In addition, congenital diseases and cancers can be induced by chromosomal rearrangements that affect the regulatory neighbourhoods of target genes^{7,8}.

With so many more potential enhancers than genes, an outstanding task is to functionally connect mammalian regulatory sequences to target genes. In this context, the three-dimensional (3D) configuration of the genome is important because it must accommodate the physical contacts between promoters and distant enhancers. Chromosome conformation studies and genetic analyses of representative loci have recently started to uncover the complex and versatile mechanisms behind target gene selection and enhancer landscape recruitment. In this Review, we discuss a few specific cases involving long-range gene regulation in mammals to illustrate emerging principles whereby remote enhancers can achieve their functions in complex genomic environments.

Evolution of mammalian enhancer landscapes

Vertebrate genomes are unique in that they contain large gene deserts with enhancers acting over distances in the megabase range (see ref. 9 for a review). Invertebrate species studied so far tend to have more local regulatory controls, which can often be recapitulated by short transgenes, such as has been shown for the roundworm *Caenorhabditis elegans*. Admittedly, in *Drosophila*, gene regulation during development is complex, with multiple enhancers acting on individual genes¹⁰ and some loci controlled by series of intricate enhancers¹¹. However, these enhancer–promoter interactions generally occur over distances shorter than 50 kb (see for example ref. 12) (Fig. 1).

The apparent restriction of megabase-sized regulations to vertebrate genomes is puzzling. It has been argued that the emergence of vertebrates was accompanied by a burst of pleiotropy and gene multi-functionality, such that crucial gene functions were co-opted for a variety of additional tasks (see refs 13 and 14 for references). This was probably achieved by multiplication of enhancers per gene of interest — a process triggered by the two genome duplications that occurred at the root of this taxon. Duplications made task-sharing among paralogous genes possible and may thus have given duplicated genes the licence to evolve additional regulations¹⁵. As a result, many vertebrate genes that are essential for important developmental pathways and active at different places and times (for example, *Hox*, *Pax*, *Fgf*, *Bmp* and *Hh*), have been kept in several paralogous copies and display complex regulatory landscapes.

Finding regulatory sequences

The complexity of the mammalian regulatory genome is revealed by the analysis of transgenic mice carrying a transposable reporter gene cassette¹⁶ used as an enhancer trap¹⁷. The staining of hundreds of embryos with an insertion at different genomic locations showed that the minimal promoter was silent at 40% of the integration sites. In nearly 60% of the embryos, however, the reporter gene was active, usually with tissue-restricted expression, showing its capacity to integrate resident regulatory signals. These specific patterns often followed that of the nearest gene. Tissue-specific transcription was also detected near housekeeping genes, suggesting that ubiquitous transcription may result from the integration of various specific cues. Although integration sites located hundreds of kilobases apart could give rise to the same expression patterns, elsewhere in

¹Hubrecht Institute-KNAW and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT, Utrecht, the Netherlands. ²School of Life Sciences, Ecole Polytechnique Fédérale, CH-1015, Lausanne, Switzerland. ³Department of Genetics and Evolution, University of Geneva, CH-1211, Geneva, Switzerland.

the genome distinct transcription patterns were identified between integration sites separated by only few kilobases, making the mouse genome a regulatory jungle¹⁶ (Fig. 2 and Box 1).

This perception is supported by the systematic mapping of enhancers, carried out either by looking at particular chromatin features associated with enhancer activity, or by multi-species alignments of non-coding DNA sequence in syntenic regions (Table 1). Although the latter approach has identified many enhancers genome wide¹⁸, enhancers may diverge in their sequences between species¹⁹. Nevertheless, DNA sequence is the prime determinant of transcription factor binding; when a human chromosome was introduced into mice, both the murine transcription factor binding profiles and the resultant gene expression patterns across the human chromosome were nearly identical when compared with human hepatocytes²⁰. Therefore, DNA motifs allowing enhancer prediction must exist and, by using experimentally derived enhancer motifs as inputs for machine-learning algorithms, several groups recently made accurate predictions^{21–23}.

Transcription-factor-binding to DNA creates chromatin signatures that can also be used to experimentally identify enhancers. Such signatures include the local opening up of chromatin (uncovered by DNaseI sensitivity^{24,25}), the presence of transcription factors and co-factors such as p300 and the deposition of histone marks such

as monomethylation of histone H3 lysine 4 (H3K4) and acetylation of H3K27, as assayed by chromatin immunoprecipitation (ChIP). Systematic mapping by these methods across cell types have uncovered millions of sites with regulatory potential, most of them recognizable only in a tissue-specific manner²⁶.

Functional screens for enhancers

Predicted enhancers should be validated in functional essays. Transgenic animals are typically used for this purpose and can validate up to 90% of the inserted DNA sequences. Recently, self-transcribing active regulatory region sequencing (STARR-seq) has facilitated screens based on enhancer activity, whereby candidate sequences stimulate their own transcription²⁷. Enhancer strength in one particular cellular context is thus reflected by the abundance of the corresponding transcripts. When applied to *Drosophila* S2 cells in which 11 million random fragments were tested for activity, a number of interesting observations were made. For example, 5% of the 5,499 regions displaying enhancer activity were bona fide transcription start sites, showing that, to some extent, promoters may also enhance transcription. Many strong enhancers were mapped near housekeeping genes, and often multiple enhancers (five or more) seemed to target the same gene, including housekeeping genes, further suggesting that ubiquitous expression may also be controlled by complex networks of tissue-specific enhancers rather than being a promoter-intrinsic feature.

When compared with DNaseI and various ChIP profiles obtained from S2 cells, a third of the enhancer sequences scored by STARR-seq were within a repressive chromatin configuration, often near silent, important developmental genes. These sequences lacked H3K27ac but carried H3K27me3 (reflecting Polycomb-mediated repression) and H3K4me1 (a modification found at enhancers). These sequences were thus recognized as functional enhancers, but they were actively silenced²⁷. This emphasizes a limitation associated with all enhancer screening methods; although a given DNA sequence may display regulatory capacity, it may not exert this property in its physiological context, raising the need not only to evaluate enhancer functionalities by genetic approaches in their natural environment, but also to study in some detail the nature of these environments. Chromosome topology has become an important parameter in this context, because it accommodates the wiring between enhancers and their target endogenous genes. Therefore, enhancer action must be considered in the 3D structure of the genome.

Enhancer action through DNA looping

The development of chromosome conformation capture (3C) technologies²⁸ has advanced our understanding of regulatory interactions in the 3D genome. 3C considers the amount of ligation products between cross-linked DNA segments as a function of their contact frequencies *in vivo*. Originally designed for the analysis of one-to-one contacts between selected pairs of genomic sequences, high-throughput variants now allow the assessment of one-to-all (4C), many-to-many (5C) or all-to-all (Hi-C) contacts, thus providing distinct details about chromatin topology²⁹. Using the β -globin (also known as *Hbb*) locus, 3C studies showed that globin genes form tissue- and differentiation-specific contacts with the distant locus control region (LCR), thus illustrating that spatial proximity to enhancers can increase promoter activity^{30,31}. Presumably, this proximity increases the local concentration of DNA binding modules, causing a nearby accumulation of cognate transcription factors at the promoter that strengthens its transcriptional output.

Transcription itself is not required for the formation of enhancer-promotor loops³²; however, the loops are required for transcription, as confirmed by enhancer-promotor loops engineered at the β -globin locus *in vivo*³³. In mouse cells that lack the erythroid-specific transcription factor Gata1, chromatin loops are absent and globin genes are silent. Although the Gata1-associated protein Ldb1

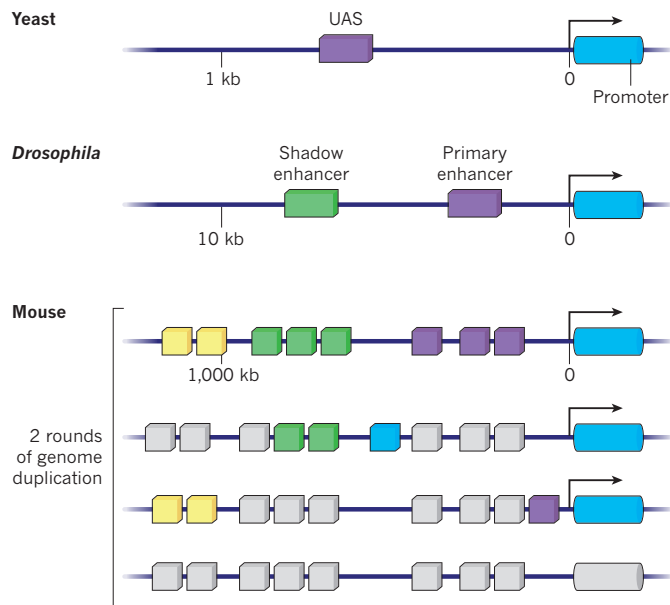


Figure 1 | Variations in long-range gene regulation. Differences exist between the enhancer–promoter distances of yeast, *Drosophila* and mice. Hypothetical and representative gene loci are shown with yeast locus in which very few enhancers, or upstream activating sequences (UAS), are usually found; those that are, exist within 1 kb (up to maximum of a few kilobases) of the promoter. In *Drosophila*, often multiple enhancers exist and they are usually located within 10 kb of the promoter. Occasionally, they are found at distances of up to 100 kb, and some complex developmental regulatory landscapes have been reported, for example the BX-C locus, which stretches over distances of around 300 kb, but such intricate regulations do not seem to be the rule in *Drosophila*. In mammals, shown here in mice, regulatory landscapes found around developmental genes often extend over several hundred kilobases to more than 1 Mb. The two rounds of genome duplication that accompanied the emergence of vertebrates may have allowed for additional regulatory complexity to develop, owing to the release of constraints associated with the target gene, thus triggering the *de novo* evolution of enhancers and the diversification of their use. As a result, different paralogous landscapes display various enhancer combinations (coloured squares). Concomitantly, large gene deserts may have evolved to prevent bystander effects and regulatory interferences.

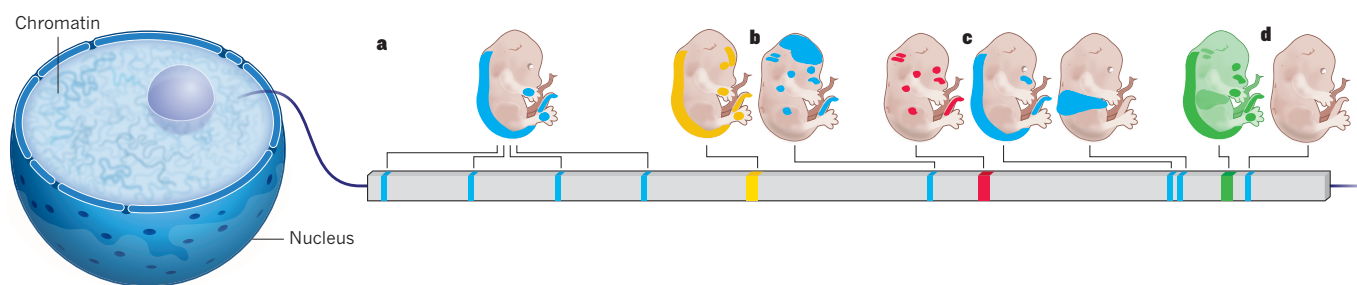


Figure 2 | The mammalian regulatory jungle. A model of three hypothetical expression patterns (yellow, red and green) and their hypothetical expression pattern at a given stage of embryonic development are shown. Embryos coloured blue show the activity of a given reporter gene integrated at different chromosomal locations (adapted from ref. 16). They illustrate that genomic context critically determines expression patterns. Thus, **a**, various insertion sites may display comparable expression patterns despite being spread over a large chromosomal interval. Note that these reporter genes incorporate most of

the regulatory activities acting on the downstream gene shown in yellow. **b**, Often, the reporter gene incorporates the enhancer activities that control the expression of one of the nearest genes (red gene). **c**, Tissue-specific reporter gene expression can sometimes be seen at sites close to housekeeping genes (green gene). In addition, two closely linked integration sites may show very distinct expression patterns that reveal highly localized regulatory circuits. **d**, At some chromosomal sites, the reporter gene is inactive and apparently not capable of capturing enhancer activity.

is no longer recruited to the β -globin gene promoter, it still binds to the LCR through other transcription factors. When artificial zinc fingers were used to force recruitment of Ldb1 to the β -globin promoter, looping with the LCR was induced in Gata1-null cells leading to robust transcription activation³³. Likewise, an isolated human LCR introduced into transgenic mice at an ectopic site was capable of *trans*-activating an endogenous β -globin gene located on another chromosome only in those cells in which an inter-chromosomal contact was established³⁴. Both studies show that contacts are necessary for enhancers to activate transcription.

Pre-established compared with *de novo* formed loops

Enhancer–promoter loops have now been identified at many loci. However, the presence of such loops is not necessarily associated with an active transcriptional outcome, suggesting that some loci display a spatial configuration, which is poised for transcription. This permissive situation is in contrast to loops that are initiated at the time of transcriptional activation because of an instructive process (Fig. 3). In such a permissive situation, a regulatory landscape exists in a preformed 3D conformation that can be used in any cell type by tissue-specific transcription factors for efficient transcription activation³⁵. In the instructive model, it is the *de novo* establishment of a chromatin configuration triggered by specific transcription factors, for example by looping, which will cause transcriptional activation. Studies on a restricted number of loci have highlighted these various alternatives.

Enhancer–promoter loops at the α -globin (also known as *Hba*) and β -globin gene loci are formed exclusively in erythroid cells, with interaction frequencies increasing during erythroid maturation^{36,37}. In addition, loop formation depends on the presence of erythroid-specific transcription factors such as Klf1 and Gata1 (refs 38, 39). Likewise, the contacts between the *SatB1* gene and its enhancer landscape found in a 800 kb large flanking gene deserts, are formed *de novo* in thymocytes in which *SatB1* is transcribed at high levels⁴⁰ (Fig. 3b). Therefore, in both cases, tissue-specific factors are necessary to trigger long-range contacts, specifically in tissues that require high transcriptional output.

By contrast, evidence for preformed configurations has been found both at the *Hox* (discussed later) and the *Shh* gene loci. *Shh* expression in posterior limb buds is necessary for correct limb development⁴¹ and is controlled by an enhancer (ZRS) located 1 Mb away, within an intron of the *Lmbr1* gene. Mutations in the ZRS cause polydactylies in both humans and mice^{42,43}. Despite this large distance, the ZRS loops and contacts the *Shh* locus⁴⁴. This loop, however, is not specific for posterior limb cells because privileged contacts are already observed in embryonic stem cells in which *Shh* is inactive⁴⁵. Also, contacts between the enhancer region and the promoter occur

even in the absence of the enhancer itself⁴⁴ and the ZRS activity can be expanded to ectopic sites in the limb bud whenever mutations recruit new sets of transcription factors⁴⁶. This suggests the existence of a preformed topology that organizes the physical proximity between the ZRS and its target *Shh* gene. In this view, the locus may be in a ‘permissive’ configuration and tissue-specific transcription factors acting through the ZRS would merely select and consolidate an existing structure. The transcription factors p53 and FOXO3 seem to act similarly through pre-existing chromatin loops^{47,48}. These two proteins are not developmental regulators, however, but factors required for cell proliferation and survival.

Preformed, permissive structures potentially offer some regulatory benefits. They may help to target enhancers to a gene of interest, thereby preventing bystander activation of unrelated neighbouring genes⁴⁹. Transcriptional activation may also be simpler to implement, because it may only involve slight and discrete variations in internal contacts within a largely conserved structure, in a way that is related to allosteric transitions of single molecules³⁵. Finally, the existence of preformed and as yet inactive regulatory landscapes could have been a rich playground for the emergence of new enhancers, because both the basic structural context and the transcriptional outcome would be available and ready to be hijacked by factors with distinct tissue specificities.

Promoter contact networks

High-throughput variants of 3C technology allow the simultaneous analysis of contacts made by multiple promoters. Comprehensive studies of such long-range promoter interactions have either used 5C to interrogate contacts across 1% of the genome⁵⁰ or the ChIA-PET approach⁵¹. ChIA-PET combines ChIP with a 3C strategy to uncover the chromatin loops formed by genomic sites that are bound by a protein of interest (in this case, RNA polymerase II)⁵¹. Both 5C and ChIA-PET studies have shown that most promoters are engaged in chromatin loops, often in cell-type-specific manners. Using 5C, contacts were found with sites showing enhancer-type chromatin signatures and with sites bound by CCCTC-binding factor (CTCF), a chromatin architecture protein known to be involved in loop formation⁵² and chromatin organization. However, most contacts were with ‘unclassified’ sites without any recognizable chromatin or sequence mark⁵⁰.

Inter-promoter contacts were also discovered by ChIA-PET and preferentially occurred between genes displaying coordinated expression. *In vitro* reporter assays further suggested that promoters can enhance each other’s activity⁵¹. Although it is often assumed that enhancers target the nearest gene, 5C data suggest that this is true in only 7% of cases. In addition, nearly 80% of long-range DNA contacts remained unaffected when intervening sequences were bound

BOX 1

Regulatory landscapes in mammalian genomes

- Around 20,000 genes
- More than 10^6 enhancers (potential regulatory sequences)
- About four enhancers contact an active gene on average per cell type⁵⁰
- Average enhancer–promoter loop size⁵⁰ is 120 kb
- Largest enhancer–promoter distance so far (SOX9, Pierre Robin disease)⁷⁴ is 1,300 kb
- 545 gene deserts (>640 kb, that is, top 3% largest deserts)⁷⁵
- Largest gene desert⁷⁵ is 5.1 Mb

by CTCF, challenging the idea that the major role of this protein is to block enhancer–promoter contacts. Finally, at least 10% of distal sites were engaged in contacts with multiple genes and, likewise, many genes formed contacts with more than one distant site⁵⁰.

The functional relevance of these complex interactions is admittedly difficult to evaluate. Enhancers defined by genetic approaches generally contact their target genes in chromosome conformation studies^{31,44} and bona fide enhancers have been isolated using 4C analyses^{35,40,53}. However, a physical contact does not necessarily reflect a functional interaction^{37,54}. Therefore, although 3C-based strategies can physically connect genes to potential enhancer sequences, genetic approaches in the appropriate *in vivo* systems are required to reveal whether and how such potential enhancers are functionally connected to a specific target gene.

Target selection by promiscuous enhancers

Mammalian promoters not only consist of a core sequence but often also contain immediate upstream binding sites for general

and tissue-specific transcription factors, which serve as a first regulatory layer to confer some tissue specificity. These additional modules can integrate regulatory activities from remote enhancers, as illustrated by the effect of a transgenic ‘orphan’ LCR (without any of the globin family of genes). When positioned into an unrelated locus, an LCR elicited a tissue-specific upregulation of many of its surrounding endogenous genes, which normally do not encounter this enhancer⁵⁵. In another study, the regulation of *Fgf8* was examined genetically. *Fgf8* encodes a protein of a pleiotropic signalling pathway and hence its transcription during development must be tightly controlled. In the 200 kb region surrounding *Fgf8*, several unrelated genes are found, as well as nearly 50 regulatory modules that bear tissue-specific information. Chromosomal rearrangements of the *Fgf8* locus *in vivo*, whereby unrelated genes were placed at the *Fgf8* position, recapitulated the highly specific developmental expression of *Fgf8* (ref. 56).

Such a prevalence of enhancer strength over promoter selectivity must lead to situations in which functionally unrelated neighbour genes share tissue-specific expression patterns. Such bystander effects do exist^{57–59}, and are best seen when genes are upregulated after genetic rearrangements. For example, the deletion of the two α -globin genes re-directs their enhancer to the *NME4* gene, 300 kb away, causing an eightfold increase in its expression⁶⁰. In addition, enhancers that control *Hoxd* genes in digits can recruit a new set of target genes after a Robertsonian translocation⁶¹. By contrast, a promoter-creating mutation in between the α -globin genes and their cognate enhancers causes the blood disorder α -thalassemia, presumably by re-allocating the enhancer away from α -globin genes⁶². In evolutionary terms, bystander effects can be prevented by restricting the number of genes within a regulatory landscape. This may explain why key developmental vertebrate genes are frequently flanked by evolutionary conserved gene deserts (or gene-poor regions) that are rich in enhancer sequences⁹. In contrast, housekeeping genes often cluster in the genome, which may help them to collectively compete for and titrate out enhancer activity, thereby buffering potential fluctuations in expression.

Table 1 | Methodology for screening for functionally relevant enhancers

Data type that identification is based on	Sensitivity	Specificity
Comparative genomics		
Screens for evolutionarily conserved sequence blocks (<i>in silico</i>)	Will not score evolutionarily diverged enhancers ¹⁹	Also identifies regulatory sites with no enhancer activity (promoters, insulators and architectural sites) and redundant enhancers, for example without target genes (orphan enhancers) ⁷⁶
ChIP-seq of transcription factors, p300, H3K27ac and H3K4me1		
Genomic screens for sites associated with protein factors and histone modifications often found at enhancers	Will not score enhancers lacking detectable signature	Also identifies non-enhancers with similar signatures and redundant enhancers, for example those without target genes (orphan enhancers)
DNaseI profiling		
Screens for genomic sites with locally opened up chromatin	Expected to identify (nearly) all sites with regulatory potential	Also identifies other regulatory sites (promoters, insulators and architectural sites) and redundant enhancers, for example those without target genes (orphan enhancers)
Enhancer trap		
Transgenic reporter genes as a read-out for local enhancer activity at different genomic locations	A medium-throughput technique: misses enhancers owing to the limited number of integration sites analysed; and enhancers incompatible with reporter gene promoter	Also identifies redundant enhancers, for example without target genes (orphan enhancers)
STARR-Seq		
Functional screens for genomic sequences across the genome for their capacity to enhance their own transcription	Will not score enhancers incompatible with reporter gene promoter	Also identifies occluded enhancers that are actively repressed in the cell of interest and redundant enhancers, for example those without target genes (orphan enhancers)
3C-based methods (promoter centred)		
Screens for chromosomal sites that physically contact a promoter of interest	Will not score infrequently contacted enhancers, enhancers located close to (<10 kb) the promoter and enhancers acting independently from promoter contact	Also identifies bystander contacts with non-enhancers and redundant enhancers

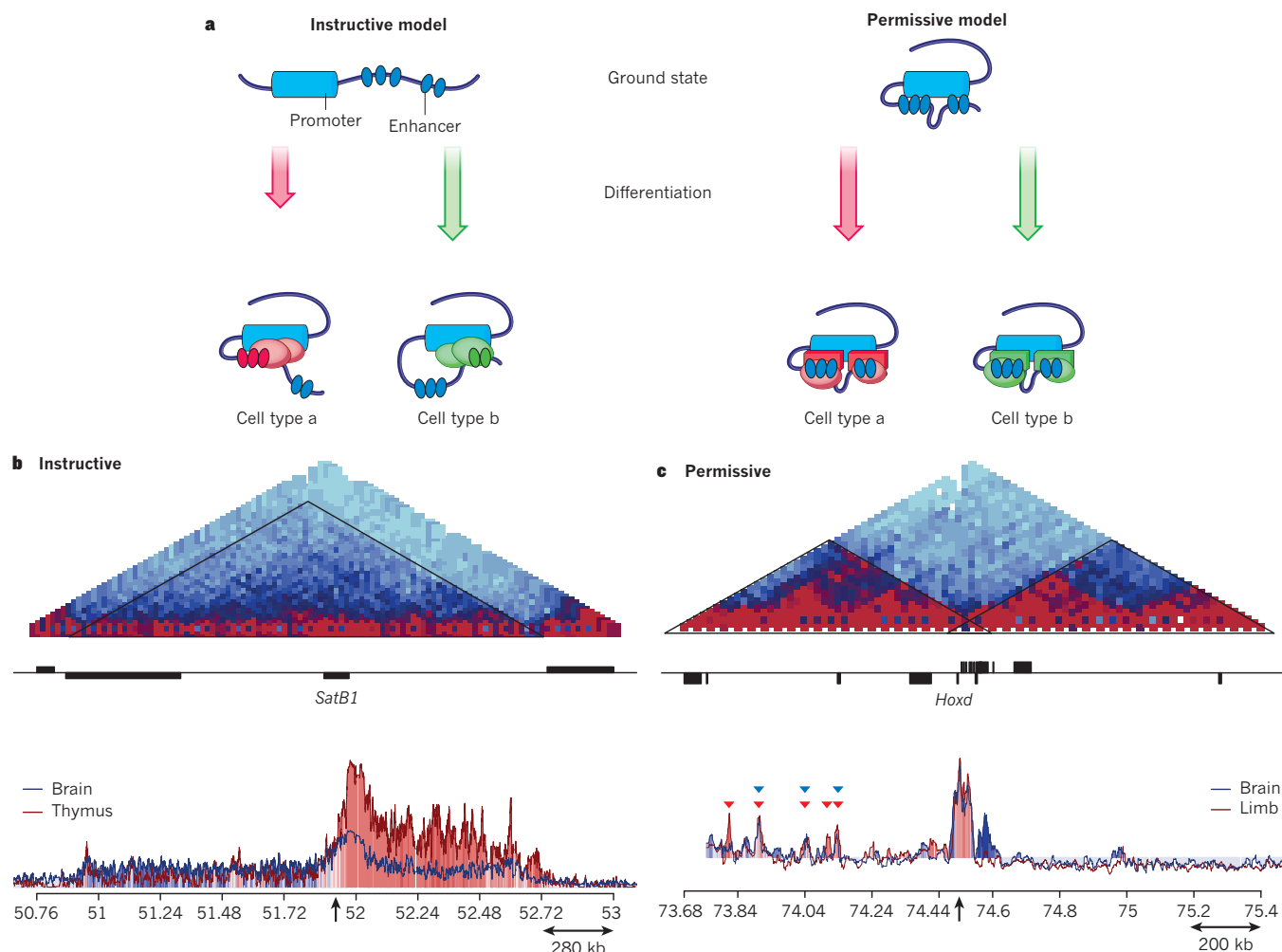


Figure 3 | Comparison of instructive and permissive model for three-dimensional controlled gene expression during differentiation. **a**, In the instructive model, tissue-specific enhancer–promoter contacts are formed *de novo* during differentiation, depending on the available transcription factors acting on the locus. In this view, no particular three-dimensional (3D) structure is formed in those cells in which the gene is inactive. In the permissive model, a preformed (ground-state) structure already exists in progenitor cells, formed either by selective promoter–enhancer interactions or by intrinsic properties of the chromatin domain. On differentiation, transcriptional activation requires merely the additional binding of tissue-specific transcription factors, which will take advantage of the configuration for immediate and robust gene activation. **b**, The *SatB1* locus is an example

of an instructive regulatory operation. TADs, identified by Hi-C (top) and site-specific contact profiles (4C data) for the *SatB1* gene promoter (arrow) (bottom). 4C reveals that robust contacts between the gene and sites across the flanking 800 kb TAD are exclusively established in thymocytes (red) that highly express *SatB1*, and not in brain cells (blue) that poorly express the gene. **c**, *Hoxd* genes are an example of a permissive regulatory operation. In limb cells (red), *Hoxd13* exclusively contacts the gene desert inside the TAD on the left, with particular interactions that involve five regulatory islands (red arrowheads). The same region, and three out of five, islands are also contacted in brain cells (blue arrowheads) in which the gene is not expressed, indicating a preformed domain that only requires subtle structural changes to support transcription in the limb (adapted from refs 35, 40, 45, 53).

Recruiting regulatory landscapes

Systematic deletions *in vivo* of potential enhancer sequences from a regulatory landscape have been done in less than a handful of studies. Although the loss of the 20 kb LCR reduced β -globin gene transcription 25- to 100-fold, the deletion of any of its five individual enhancer modules only reduced expression by a factor of 1.03–1.7 (ref. 63). These individual sites thus seem to collaborate, probably by aggregating into a single active chromatin hub³¹, which would interact with one target gene at a time⁶⁴. At the α -globin gene cluster a single enhancer, located 35 kb away in the intron of a housekeeping gene, accounts for 95% of the elevated α -globin transcript levels. Three other enhancers do form physical contacts but seem to be genetically redundant. Only when they are deleted along with the main enhancer do they abolish the remaining 5% of transcription³⁷.

Multiple enhancers have also been described in gene deserts adjacent to the *HoxD* gene cluster. *Hox* genes encode transcription factors that have key roles during the patterning of the various body axes in

bilateral animals. In vertebrates, long-range enhancers have evolved within the flanking gene deserts to accompany the emergence of vertebrate-specific features such as the appendicular skeleton. Analyses of deletions *in vivo* and 4C experiments have revealed that the digit enhancers in the centromeric desert form a regulatory archipelago — a set of islands with contacts between themselves and with the target genes. These regulatory islands complement each other to reach the final transcriptional outcome, both in the quantity of transcripts and in their spatial distribution. Some — but not all — of these contacts are maintained when the enhancers are inactive and hence they form a permissive background configuration^{35,53} (Fig. 3c). A comparable situation was observed at the opposite telomeric gene desert, in which several enhancers scattered throughout 1 Mb of DNA regulate their target *Hoxd* genes in the developing forearm. Therefore, *Hoxd* genes physically recruit different and preformed regulatory landscapes at different times and in different cells, initially to allow forearm construction and subsequently, to help form digits³⁵.

TADs as moulds for enhancer–promoter contacts

How are regulatory contacts coordinated in the 3D genome and what are the length scales over which enhancers can find their targets? Recent reports indicate that the regulatory landscapes mapped around the *Hoxd* and *Shh* loci match well with so-called topologically associating domains (TADs). TADs^{45,65,66} were defined as chromosomal regions within which sequences preferentially contact each other, based on genome-wide interaction maps generated by Hi-C⁶⁷. These domains, which are conserved among mammalian species but not found in yeast, are about 1 Mb and are separated by boundary regions that often contain CTCF-binding sites, housekeeping genes, transfer RNA genes or short interspersed elements^{45,65,66}.

The *Shh* gene for instance falls into a TAD, which precisely spans the region extending from the gene to its most distal ZRS enhancer. ZRS–promoter contacts are therefore contained within, or secured by, the overall 3D shape of the chromosomal domain. Likewise, both regulatory landscapes flanking the *Hoxd* gene cluster exactly match two topological domains. In this case, the *Hox* gene cluster itself lies between these two TADs⁴⁵, and the genes located right at the boundary are capable of switching their contacts from one TAD to the other³⁵, suggesting that enhancers acting over a given gene or set of genes may not always be restricted to a single TAD. In this view, TADs may sometimes also be used as large units of tissue-specific transcription.

Are TADs a cause or an effect?

TADs are, for the most part, already formed in embryonic stem cells and therefore seem to exist regardless of the transcription status^{45,49}. However, recent topology maps at higher resolution revealed extensive structural reorganization at the sub-megabase scale during differentiation⁶⁸, with distinct hierarchical roles for different architectural proteins. At the 0.1 to 1 Mb scale, CTCF and cohesin may anchor constitutive interactions around development-specific genes. Below 100 kb, mediator, a protein complex conventionally associated with enhancer activity, and cohesin might cooperate to bridge tissue-specific enhancer–promoter interactions⁶⁸.

Topological domains may reflect an “inherent property of the mammalian genome”⁴⁵ that could help to limit the distance over which enhancers operate, providing a mould for enhancer–promoter interactions to occur and thus limiting undesired bystander effects⁴⁹. Alternatively, TADs might merely result from pervasive enhancer–promoter interactions and thus illustrate the existence of a pre-regulatory genome, comprised of poised structures and lacking the final tissue-specific factor to become active. Genetic approaches to perturb TADs by means of chromosomal rearrangements^{35,65} will help us both to understand the mechanisms underlying TAD formation and discriminate between these hypotheses.

Very long-range regulation

So far, little evidence exists for mammalian enhancers to act beyond the few megabase scale (that is, more than the TAD organizational level). Although circumstantial evidence supports inter-chromosomal gene regulation⁶⁹, it has rarely been verified genetically and whenever such observations were complemented by the deletion of enhancers *in vivo*, only neighbouring genes located on the same chromosome were affected⁷⁰. Mammalian *trans*-activation was, however, seen in transgenic mice carrying an orphan LCR, which was able to activate its natural β -globin target gene on another chromosome. This nevertheless happened through fortuitous inter-chromosomal contacts, made in only a subset of cells, which consequently displayed elevated levels of globin expression³⁴. Although artificial, this experiment shows what could probably be expected from a genome that, beyond the level of TADs, is structured in a probabilistic manner with the overall shape and relative location of chromosomes being different from cell to cell: productive inter-chromosomal enhancer–promoter interactions may exist but are

likely to result in variegated expression. For pan-cellular expression control, enhancers therefore seem best positioned in *cis*, within domains of preferred chromosomal contacts, such as TADs.

Perspectives

High-throughput technologies for the identification of regulatory sequences have provided us with a wealth of information about the regulatory potential of our genome. The difficult task ahead will be to functionally connect genes to regulatory sequences and establish the relevance of their interactions. The systematic application of functional screening methods such as STARR-Seq²⁷ to a wide range of mammalian cell types and tissues may thus help us to distinguish between the millions of sites that have been identified so far by sequence- and chromatin-based enhancer screening methods. The enhancer classifications currently used (for example, ‘poised’ when displaying H3K4me1 only, or ‘active’ when showing both H3K4me1 and H3K27ac marks) are admittedly incomplete and will have to be refined based on their regulatory capacities. Are enhancers strictly tissue-specific or are they sequences carrying tissue-invariant enhancer activity? What causes the occlusion⁷¹, in a natural context, of regulatory sites that harbour STARR activity in the same cell type?

In addition, although a given DNA sequence may be either suspected (by DNaseI profiling or ChIP-seq), or shown (by STARR-Seq), to have a regulatory capacity, its physiological relevance will have to be established in the appropriate developmental context. As a prerequisite, high resolution contact maps based on ultra-deep sequencing of Hi-C data can be expected for all relevant developmental cell types, which will allow us to physically connect potential regulatory sequences to target genes. Such detailed contact maps will help us to uncover the hierarchical folding principles of our chromosomes and clarify the degree of developmental conservation of structural domains like TADs and sub-TADs across the entire genome. It may also provide a topological framework for long-range enhancer–promoter contacts and might even inform us about those constraints underlying inter-species syntenic DNA segment conservation (see ref. 49). Although current Hi-C and 4C strategies are well equipped to perform these tasks, improvements will be necessary both to computational analysis tools and to downscaling the experimental material, whenever scarce cellular populations will be considered in the small developing mammalian embryo.

Once a particular DNA sequence is found in an open, nucleosome-depleted configuration, bound by tissue-specific transcription factors, displaying STARR-Seq activity and capable of establishing physical contacts with an endogenous gene, the question still remains as to whether this site controls the developmental expression of the target gene. Genetic approaches whereby all the parameters listed above can be perturbed in an ontogenic context will be necessary. Only a few complex regulatory landscapes have been dissected *in situ* so far and they have shown different aspects of enhancers with different recruitment strategies, different means of target gene selection and different ways of cooperating with other regulatory sites. The application of new strategies for site-directed genome editing^{72,73} may help to clarify how regulatory sites and genes functionally orchestrate developmental gene expression programs and, consequently, how failures in regulatory wiring can cause diseases. However, genetics analysis on its own will not easily answer these questions because many components of developmental pathways are notoriously redundant. In addition, the developing mammalian embryo is very efficient at implementing compensatory mechanisms, whenever sequence modifications are induced. As a result, current analytical tools will have to be streamlined such that the impact of a genetic manipulation can be investigated as exhaustively and objectively as possible. ■

Received 2 June; accepted 2 September 2013.

1. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
 2. Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**, 327–339 (2011).
 3. Spitz, F. & Furlong, E. E. Transcription factors: from enhancer binding to developmental control. *Nature Rev. Genet.* **13**, 613–626 (2012).
 4. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
 5. Stamatiouyannopoulos, J. A. What does our genome encode? *Genome Res.* **22**, 1602–1611 (2012).
 6. Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
 7. Kleinjan, D. A. & van Heyningen, V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am. J. Hum. Genet.* **76**, 8–32 (2005).
 8. Weischenfeldt, J., Symmons, O., Spitz, F. & Korbel, J. O. Phenotypic impact of genomic structural variation: insights from and for human disease. *Nature Rev. Genet.* **14**, 125–138 (2013).
 9. Montavon, T. & Duboule, D. Landscapes and archipelagos: spatial organization of gene regulation in vertebrates. *Trends Cell Biol.* **22**, 347–354 (2012).
 10. Lagha, M., Bothma, J. P. & Levine, M. Mechanisms of transcriptional precision in animal development. *Trends Genet.* **28**, 409–416 (2012).
 11. Maeda, R. K. & Karch, F. Gene expression in time and space: additive vs hierarchical organization of *cis*-regulatory regions. *Curr. Opin. Genet. Dev.* **21**, 187–193 (2011).
 12. Levine, M. Transcriptional enhancers in animal development and evolution. *Curr. Biol.* **20**, R754–R763 (2010).
 13. Duboule, D. & Wilkins, A. S. The evolution of 'bricolage'. *Trends Genet.* **14**, 54–59 (1998).
 14. Kirschner, M. & Gerhart, J. Evolvability. *Proc. Natl Acad. Sci. USA* **95**, 8420–8427 (1998).
 15. Ohno, S. *Evolution by Gene Duplication*. (Springer, 1970).
 16. Ruf, S. *et al.* Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nature Genet.* **43**, 379–386 (2011).
- In this paper, the authors use a transposable reporter gene system in mice to probe for enhancer activity *in vivo* and show widely varying reporter expression patterns at hundreds of genomic integration sites.**
17. O'Kane, C. J. & Gehring, W. J. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl Acad. Sci. USA* **84**, 9123–9127 (1987).
 18. Pennacchio, L. A. *et al.* *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502 (2006).
 19. Odom, D. T. *et al.* Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nature Genet.* **39**, 730–732 (2007).
 20. Wilson, M. D. *et al.* Species-specific transcription in mice carrying human chromosome 21. *Science* **322**, 434–438 (2008).
 21. Zinzen, R. P., Girardot, C., Gagneur, J., Braun, M. & Furlong, E. E. Combinatorial binding predicts spatio-temporal *cis*-regulatory activity. *Nature* **462**, 65–70 (2009).
 22. Lee, D., Karchin, R. & Beer, M. A. Discriminative prediction of mammalian enhancers from DNA sequence. *Genome Res.* **21**, 2167–2180 (2011).
 23. Gorkin, D. U. *et al.* Integration of ChIP-seq and machine learning reveals enhancers and a predictive regulatory sequence vocabulary in melanocytes. *Genome Res.* **22**, 2290–2301 (2012).
 24. Crawford, G. E. *et al.* DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nature Methods* **3**, 503–509 (2006).
 25. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
 26. Hardison, R. C. & Taylor, J. Genomic approaches towards finding *cis*-regulatory modules in animals. *Nature Rev. Genet.* **13**, 469–483 (2012).
 27. Arnold, C. D. *et al.* Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074–1077 (2013).
- In this paper, the authors report a genome-wide screening method that quantifies the self-transcription of random DNA segments to identify sequence modules with enhancer activity.**
28. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
 29. de Wit, E. & de Laat, W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* **26**, 11–24 (2012).
 30. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. & Fraser, P. Long-range chromatin regulatory interactions *in vivo*. *Nature Genet.* **32**, 623–626 (2002).
 31. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
 32. Palstra, R. J. *et al.* Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE* **3**, e1661 (2008).
 33. Deng, W. *et al.* Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
- In this study, it is shown that the artificial physical tethering of a distant enhancer to a gene promoter leads to highly activated transcription, illustrating that enhancer contacts causally underlie transcriptional output.**
34. Noordermeer, D. *et al.* Variegated gene expression caused by cell-specific long-range DNA interactions. *Nature Cell Biol.* **13**, 944–951 (2011).
- This article shows that inter-chromosomal gene activation by an ectopic enhancer leads to variegated gene expression in subsets of mouse cells.**
35. Andrey, G. *et al.* A switch between topological domains underlies *HoxD* genes collinearity in mouse limbs. *Science* **340**, 1234167 (2013).
 36. Palstra, R. J. *et al.* The β -globin nuclear compartment in development and erythroid differentiation. *Nature Genet.* **35**, 190–194 (2003).
 37. Vernimmen, D., De Gobbi, M., Sloane-Stanley, J. A., Wood, W. G. & Higgs, D. R. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041–2051 (2007).
 38. Drissen, R. *et al.* The active spatial organization of the β -globin locus requires the transcription factor EKLF. *Genes Dev.* **18**, 2485–2490 (2004).
 39. Vakoc, C. R. *et al.* Proximity among distant regulatory elements at the β -globin locus requires GATA-1 and FOG-1. *Mol. Cell* **17**, 453–462 (2004.12.028 (2005).
 40. van de Werken, H. J. *et al.* Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nature Methods* **9**, 969–972 (2012).
 41. Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
 42. Lettice, L. A. *et al.* A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).
 43. Sagai, T., Hosoya, M., Mizushima, Y., Tamura, M. & Shiroishi, T. Elimination of a long-range *cis*-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803 (2005).
 44. Amano, T. *et al.* Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47–57 (2009).
 45. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
 46. Lettice, L. A. *et al.* Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Dev. Cell* **22**, 459–467 (2012).
 47. Eijkelenboom, A. *et al.* Genome-wide analysis of FOXO3 mediated transcription regulation through RNA polymerase II profiling. *Mol. Syst. Biol.* **9**, 638 (2013).
 48. Melo, C. A. *et al.* eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell* **49**, 524–535 (2012).
 49. Nora, E. P., Dekker, J. & Heard, E. Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? *BioEssays* **5**, 201300040 (2013).
 50. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113 (2012).
- This report is the first systematic analysis, based on 5C, of promoter-centred DNA contacts, revealing complex and dynamic contact networks, which change during development.**
51. Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012).
 52. Splinter, E. *et al.* CTCF mediates long-range chromatin looping and local histone modification in the β -globin locus. *Genes Dev.* **20**, 2349–2354 (2006).
 53. Montavon, T. *et al.* A regulatory archipelago controls *Hox* genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
- In this paper, the authors report the existence of a regulatory archipelago that controls transcription in developing digits and contains several islands with various quantitative and qualitative contributions.**
54. Bender, M. A. *et al.* Flanking HS-62.5 and 3-HS1, and regions upstream of the LCR, are not required for β -globin transcription. *Blood* **108**, 1395–1401 (2006).
 55. Noordermeer, D. *et al.* Transcription and chromatin organization of a housekeeping gene cluster containing an integrated β -globin locus control region. *PLoS Genet.* **4**, e1000016 (2008).
 56. Marinić, M., Aktas, T., Ruf, S. & Spitz, F. An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530–542 (2013).
 57. Caijiao, I., Zhang, A., Yoo, E. J., Cooke, N. E. & Liebhauer, S. A. Bystander gene activation by a locus control region. *EMBO J.* **23**, 3854–3863 (2004).
 58. Spitz, F., Gonzalez, F. & Duboule, D. A global control region defines a chromosomal regulatory landscape containing the *HoxD* cluster. *Cell* **113**, 405–417 (2003).
 59. Zuniga, A. *et al.* Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression. *Genes Dev.* **18**, 1553–1564 (2004).
 60. Lower, K. M. *et al.* Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. *Proc. Natl Acad. Sci. USA* **106**, 21771–21776 (2009).
 61. Tschopp, P., Fraudeau, N., Bena, F. & Duboule, D. Reshuffling genomic landscapes to study the regulatory evolution of Hox gene clusters. *Proc. Natl Acad. Sci. USA* **108**, 10632–10637 (2011).
 62. De Gobbi, M. *et al.* A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* **312**, 1215–1217 (2006).
 63. Bender, M. A. *et al.* Targeted deletion of 5-HS1 and 5-HS4 of the β -globin locus control region reveals additive activity of the DNaseI hypersensitive sites. *Blood* **98**, 2022–2027 (2001).
 64. Wijgerde, M., Grosveld, F. & Fraser, P. Transcription complex stability and chromatin dynamics *in vivo*. *Nature* **377**, 209–213 (1995).
 65. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
 66. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
- Refs 45, 65 and 66 report that mammalian and fly chromosomes are structurally subdivided into topological domains with flanking boundaries that hamper DNA contacts across.**

67. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
68. Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
69. Williams, A., Spilianakis, C. G. & Flavell, R. A. Interchromosomal association and gene regulation in trans. *Trends Genet.* **26**, 188–197 (2010).
70. Fuss, S. H., Omura, M. & Mombaerts, P. Local and *cis* effects of the H element on expression of odorant receptor genes in mouse. *Cell* **130**, 373–384 (2007).
71. Gaetz, J. *et al.* Evidence for a critical role of gene occlusion in cell fate restriction. *Cell Res.* **22**, 848–858 (2012).
72. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nature Biotechnol.* **29**, 143–148 (2011).
73. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
74. Kleinjan, D. J. & Coutinho, P. *Cis*-ruption mechanisms: disruption of *cis*-regulatory control as a cause of human genetic disease. *Brief. Funct. Genomic Proteomic* **8**, 317–332 (2009).
75. Ovcharenko, I. *et al.* Evolution and functional classification of vertebrate gene deserts. *Genome Res.* **15**, 137–145 (2005).
76. Ahituv, N. *et al.* Deletion of ultraconserved elements yields viable mice. *PLoS Biol.* **5**, e234 (2007).

Acknowledgements The authors thank Y. Oz for preparing the figures, and G. Geeven, M. Janssen, G. Andrey, L. Beccari, T. Montavon and M. Leleu for their help. The authors' laboratories are supported by funds from the Dutch Scientific Organization (NWO) (714012002 and 724012003 (VICI)), EU grant 2010-259743 (MODHEP), a KWF Dutch Cancer Foundation grant (2009-4459) and a NanoNextNL grant (to W.dL) and the Ecole Polytechnique Fédérale de Lausanne, the University of Geneva, the Swiss National Research Foundation, the ERC grant SystemsHox.ch and the FP7 program IDEAL (to D.D.).

Author Information Reprints and permissions information is available at www.nature.com/reprint. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at go.nature.com/cwhy4j. Correspondence should be addressed to W.dL (w.delaat@hubrecht.eu) or D.D. (denis.duboule@epfl.ch).

Odour receptors and neurons for DEET and new insect repellents

Pinky Kain^{1*}, Sean Michael Boyle^{2*}, Sana Khalid Tharadra¹, Tom Guda¹, Christine Pham³, Anupama Dahanukar^{1,2,3,4} & Anandasankar Ray^{1,2,3,4}

There are major impediments to finding improved DEET alternatives because the receptors causing olfactory repellency are unknown, and new chemicals require exorbitant costs to determine safety for human use. Here we identify DEET-sensitive neurons in a pit-like structure in the *Drosophila melanogaster* antenna called the sacculus. They express a highly conserved receptor, *Ir40a*, and flies in which these neurons are silenced or *Ir40a* is knocked down lose avoidance to DEET. We used a computational structure–activity screen of >400,000 compounds that identified >100 natural compounds as candidate repellents. We tested several and found that most activate *Ir40a*⁺ neurons and are repellents for *Drosophila*. These compounds are also strong repellents for mosquitoes. The candidates contain chemicals that do not dissolve plastic, are affordable and smell mildly like grapes, with three considered safe in human foods. Our findings pave the way to discover new generations of repellents that will help fight deadly insect-borne diseases worldwide.

Blood-feeding insects transmit deadly diseases such as malaria, dengue, lymphatic filariasis and West Nile fever to hundreds of millions of people, causing immense suffering and more than a million deaths every year. Insect repellents can be very effective in reducing disease transmission by blocking contact between blood-seeking insects and humans.

N,N-diethyl-*meta*-toluamide (DEET) has remained the primary insect repellent used for more than 60 years. However, DEET has little effect on disease control in endemic regions due to high costs and inconvenience of continuous application on skin at high concentrations. DEET also dissolves some plastics, synthetic fabrics and painted surfaces¹. Additionally, DEET inhibits mammalian acetylcholinesterase². Instances of DEET resistance have also been reported in flies³ and mosquitoes^{4,5}. However, the main barriers in developing improved repellents are the estimated cost for identification⁶ and the subsequent cost of safety analyses for new chemistries.

A significant challenge in finding improved DEET substitutes is that the target receptors through which it repels insects are unknown. Recent studies have given rise to many different models of DEET action. Pure DEET causes inhibition^{7,8} or mild electrophysiological modification of neural responses to weakly-activating odours in *Drosophila* antennal olfactory neurons⁹, but whether these effects contribute to repellency is unknown. Mosquitoes can also directly detect DEET¹⁰ and mutations in the *orco* co-receptor gene in *Aedes aegypti* cause reduction in repellency¹¹. Some DEET-sensitive olfactory neurons have been identified in *Culex quinquefasciatus*¹⁰ and *A. aegypti*⁵, but it is not yet known whether they are responsible for repellency or which odour receptors they express. A broadly tuned larval odour receptor responds to DEET^{12,13}; however, its role in avoidance in larval or adult mosquitoes has not been demonstrated. Not only can more than one pathway contribute to olfactory repellency, analyses are further confounded by the observation that DEET also activates bitter taste neurons that mediate contact-avoidance in *Drosophila*^{14,15}.

DEET is detected by neurons of the sacculus

To identify the elusive DEET-sensing neurons of the olfactory system in an unbiased manner, we used the nuclear factor of activated T cells (NFAT)-based system to report DEET-evoked neural activity through expression of green fluorescent protein (GFP) in *Drosophila melanogaster*¹⁶ (Fig. 1a). Exposure to 10% DEET resulted in an increase in expression of GFP in neurons that innervate sensilla within the sacculus, a pit-like structure in the antenna (Fig. 1b, c, Supplementary Fig. 1a and Supplementary Video 1). The dendrites of GFP⁺ neurons primarily innervated the most distal chamber (I) of the sacculus (Fig. 1c and Supplementary Fig. 1b). Previous studies of DEET overlooked the sacculus because it is intractable to traditional electrophysiology methods.

Contrary to expectations from a previous report¹⁷, we were unable to find DEET-activated reporter expression in odorant receptor neurons (ORNs) of the maxillary palps (Fig. 1b). We therefore performed single-sensillum electrophysiology analyses and found that the previously reported *Or42a*⁺ pb1A neurons responded poorly to DEET, but strongly to hexane that was used as solvent in the previous study (Supplementary Fig. 2a, b).

ORNs innervating the sacculus do not express *Or* genes, but instead members of a conserved ionotropic receptor (*IR*) gene family^{18–21}. In the antennal lobes robust DEET-dependent GFP was detected in the characteristic ‘column’ glomerulus (Fig. 1d and Supplementary Fig. 3a), which is innervated by axons of *Ir40a*-expressing neurons of the sacculus¹⁸. Faint GFP was also observed in the *Or67d*⁺ DA1 glomerulus, which is probably caused by exposure to male pheromone *cis*-vaccenyl acetate (cVA) in the assay, because the cVA-responsive *Or67d*⁺ at1 neuron did not respond to DEET (Supplementary Fig. 2c). The DC4 glomerulus, which is innervated by other sacculus ORNs that express *Ir64a*¹⁹, showed a very faint signal as well (Supplementary Fig. 3a). The simplest interpretation of these results is that *Ir40a*⁺ sacculus ORNs innervating chamber I and projecting to the column glomerulus may represent a chief olfactory detection pathway for DEET.

Consistent with previous electrophysiological analyses^{14,15}, we found DEET-dependent GFP expression in gustatory neurons of the labellum

¹Department of Entomology, University of California, Riverside, California 92521, USA. ²Genetics, Genomics and Bioinformatics Program, University of California, Riverside, California 92521, USA.

³Neuroscience Program, University of California, Riverside, California 92521, USA. ⁴Institute of Integrative Genome Biology, University of California, Riverside, California 92521, USA.

*These authors contributed equally to this work.

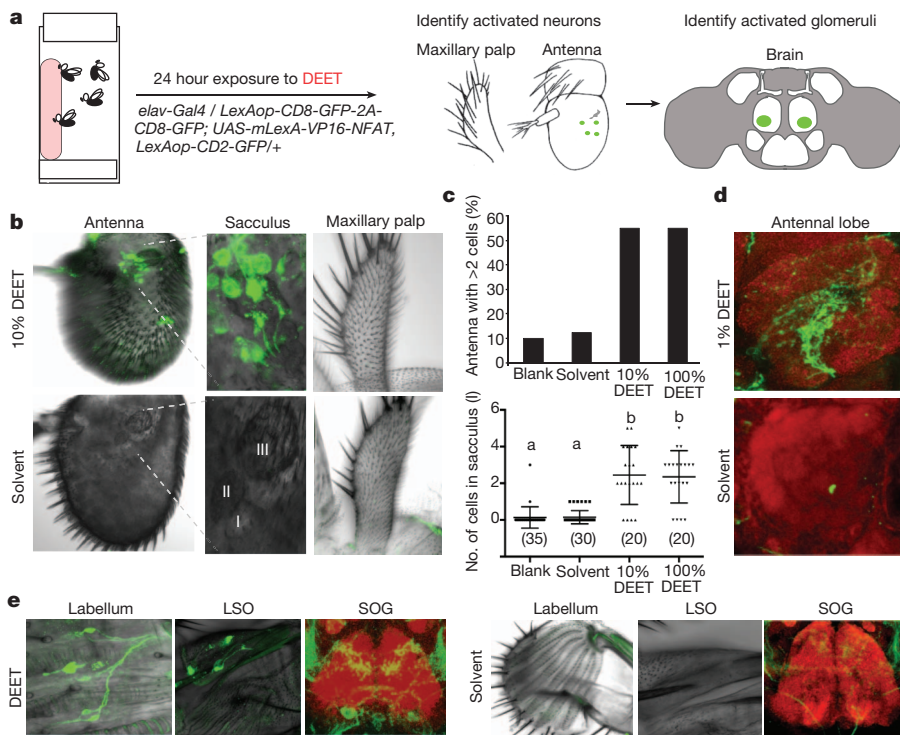


Figure 1 | DEET is detected by *Ir40a*⁺ sacculus neurons. **a**, Schematic of the NFAT (CaLexA)-based method to label neurons activated by DEET. **b**, Confocal micrographs of olfactory organs from flies stimulated with 10% DEET or solvent (acetone). **c**, Quantification of GFP⁺ antennae (upper graph) and mean numbers of GFP⁺ cells in chamber I (lower graph). **d**, GFP⁺ axonal termini in antennal lobes of flies treated as indicated. **e, f**, Expression of GFP in the labellum, labral sense organ (LSO) and subesophageal ganglion (SOG). Anti-GFP (green) and anti-nc82 (red). For SOG, dorsal is top. Error bars indicate s.e.m.

(Fig. 1e). In addition, we observed DEET-dependent GFP in neurons innervating the labral sense organ (LSO) of the pharynx (Fig. 1e). The DEET activity mapped to neurons marked by *Gr33a* and *Gr89a*, which are bitter-sensing deterrent neurons (Supplementary Fig. 3b). Axonal projections of DEET-sensitive gustatory neurons in the subesophageal ganglion (SOG) revealed arborization patterns similar to those of taste neurons originating in the labellum and the pharynx (Fig. 1e, Supplementary Fig. 3b).

In order to directly test physiological responses of the sacculus *Ir40a*⁺ ORNs to DEET we performed *in vivo* calcium imaging in flies expressing GCaMP3 using *Ir40a-Gal4*^{18,22,23}. *Ir40a* neurons show robust activation in response to a puff of DEET delivered from an atomizer but not to control solvent dimethylsulphoxide (DMSO) (Fig. 2a, b). Moreover, the DEET response is dependent on *Ir40a* (Fig. 2c).

In order to test whether the *Ir40a*⁺ ORNs are required for DEET repellency we blocked synaptic transmission in these neurons using *Ir40a-Gal4* to express the active form of tetanus toxin (TNTG)²⁴. We used a trap lured by 10% apple cider vinegar (ACV) in which a DEET-treated filter paper was placed inside the trap. Avoidance was significantly decreased in *Ir40a-TNTG* flies as compared to various controls, including a non-functional version of the tetanus toxin (IMPTV), suggesting that *Ir40a*⁺ neurons are required for DEET repellency (Fig. 2d). All genotypes exhibited attraction to 10% ACV in two-choice trap assays (Supplementary Fig. 4a).

Ir40a is necessary for DEET avoidance

To test directly whether *Ir40a* is required for olfactory avoidance to DEET, we examined the behaviour of flies in which *Ir40a* was knocked down pan-neuronally using an *elav-Gal4* driver to express a *UAS-Ir40a* RNA interference (RNAi) construct. In two-choice trap assays (Fig. 3a), we found a significant loss of DEET avoidance in the *Ir40a* RNAi flies compared to control flies (Fig. 3b). Similar results were obtained when *Ir40a* RNAi was executed selectively in *Ir40a*⁺ ORNs using two independent *UAS-Ir40a* RNAi transgenes (Fig. 3c). Not only was avoidance completely abolished, *Ir40a* knockdown flies actually showed a mild attraction to the DEET trap. Attraction to ACV was unaffected (Supplementary Fig. 4b, c).

We next wanted to rule out the possibility of a developmental role for *Ir40a*. We therefore suppressed expression of *Ir40a*-RNAi during

development using a temperature-sensitive *Gal80^{ts}* transgene (Fig. 3d). Flies were raised at the permissive temperature (18 °C) until just before adult eclosion, at which point they were left at 18 °C (RNAi off) or shifted to the *Gal80^{ts}* restrictive temperature 29 °C (RNAi on). Behavioural assays performed four days after the temperature shift showed that *Ir40a* RNAi in the adult was sufficient to abolish DEET avoidance when RNAi was induced in *Ir40a*⁺ ORNs (Fig. 3e, knockdown).

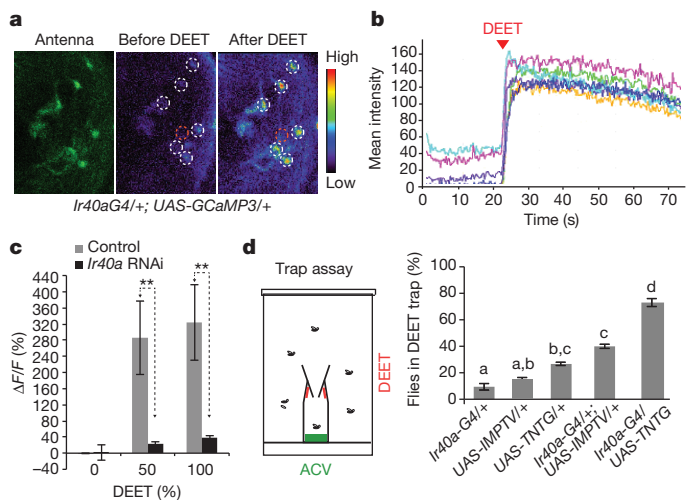


Figure 2 | *Ir40a* neurons detect DEET and are required for repellency.

a, Images of calcium activity in *Ir40a-Gal4/+;UAS-GCaMP3/+* neurons colour-coded as indicated (right). Measurements taken from areas in dashed circles: cells (white), background (red). **b**, Mean fluorescence intensities for six different cells. Red arrowhead indicates onset of ~2-s puff of DEET. **c**, Mean percentage change in fluorescence intensity after application of ~2-s indicated stimulus; genotypes were *Ir40a-Gal4/+;UAS-GCaMP3/+* (control) and *Ir40a-Gal4/Ir40a-Gal4;UAS-GCaMP3/UAS-Ir40a* RNAi (line number 2) indicated as (*Ir40a*-RNAi). **d**, Schematic (left) and results (right) for DEET-treated trap assays for indicated genotypes. Letters indicate statistical significance, $P \leq 0.008$, one-way ANOVA with Tukey's post hoc analysis. Error bars represent s.e.m.

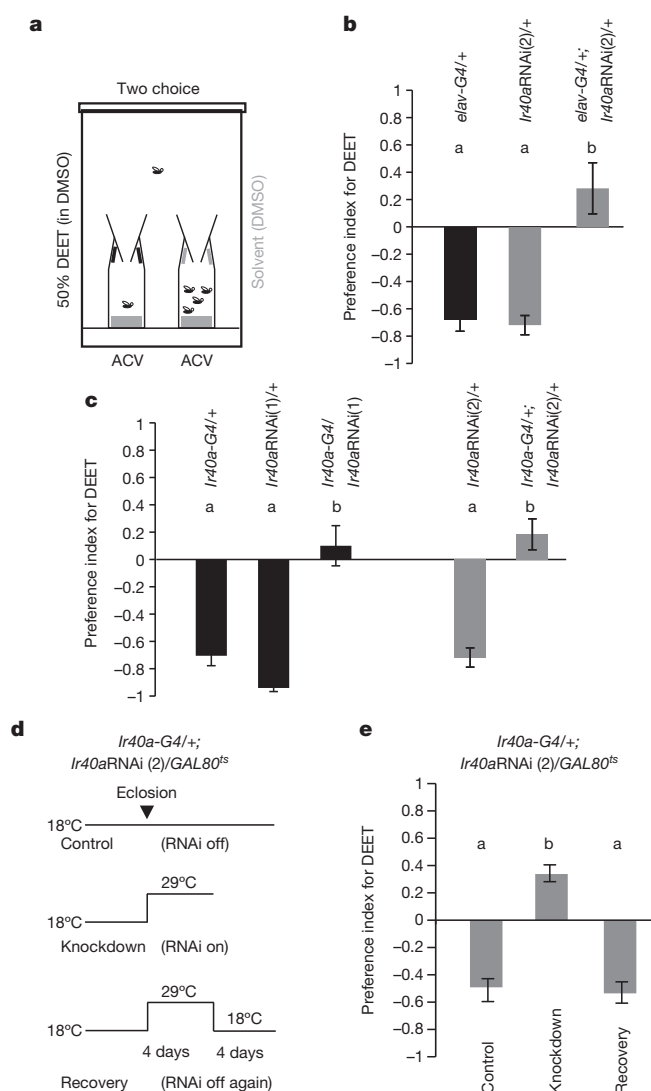


Figure 3 | *Ir40a* is required for DEET avoidance. **a**, Set-up for behavioural two-choice assay. **b**, **c**, Mean preference index of indicated genotypes for DEET in two-choice assays using *elav-Gal4* (**b**) and *Ir40a-Gal4* (**c**). $n = 6$ trials (20 flies per trial) except *elav-Gal4/+; Ir40aRNAi(2)* $n = 10$ trials and RNAi experiments with *Ir40a-Gal4* $n = 12$ trials each. **d**, Genotype and schematic for post-developmental knockdown and recovery of *Ir40a*. **e**, Mean DEET preference index of flies derived from indicated treatments in two-choice assays. $n = 6$ trials for all conditions, with 20 flies per trial. For **b–e**, $P < 0.001$, one-way ANOVA with Tukey's post hoc analysis. Error bars represent s.e.m.

Moreover, DEET avoidance was completely restored when flies were returned to the *Gal80^{ts}* permissive temperature (Fig. 3e, recovery). Attraction to ACV was unaffected (Supplementary Fig. 4d). Taken together, these experiments demonstrate that *Ir40a* is required in adult *Ir40a⁺* sacculus ORNs for olfactory avoidance of DEET.

In silico prediction of new repellents

Identification of DEET receptors and neurons offers a powerful system to screen for improved repellents. However, volatile chemical space that can be exploited to find DEET substitutes is vast and therefore poses unfeasible requirements in terms of cost and time to screen. The receptor structure is unavailable for screening and the most effective repellents may require detection by both olfactory and gustatory pathways. To circumvent these limitations we developed a high-throughput chemical informatics screen. Previous studies using such structure-activity approaches have given encouraging results²⁵.

We identified structural features shared by DEET and other known repellents and used them to screen a vast library of compounds *in silico*

for the presence of these features. We assembled a training set of known repellents that included: the two commercially approved repellents DEET and picaridin; 34 *N*-acyl piperidines²⁵ that were identified by structural relatedness to picaridin; natural repellents eucalyptol, linalool, alpha-thujone and beta-thujone^{10,26,27}; and a structurally diverse panel of other odours as negatives^{28,29}. We focused on a descriptor-based computational approach and using a sequential-forward-selection method³⁰ we incrementally identified a unique subset of 18 descriptors that were highly correlated with repellency (correlation of 0.912) (Fig. 4a and Supplementary Table 1). The repellents clustered together if the optimized descriptor subset was used to calculate Euclidean distances amongst odorants of the training set (Fig. 4b).

The optimized descriptor set was used to train a support vector machine (SVM), which is a well-known supervised learning approach³¹, to predict compounds that shared optimized structural features with known repellents (Fig. 4a). A fivefold cross-validation on the training set of repellents was performed and a mean receiver-operating-characteristic (ROC) analysis curve generated. The area under curve (AUC) was determined to be high (0.994), indicating that the *in-silico* approach was extremely effective at predicting repellents from compounds that were excluded from the training set (Fig. 4c).

We next used the 18-optimized-descriptor and SVM method to screen *in silico* a large virtual chemical library consisting of >440,000 volatile-like chemicals. Inspection of the top 1,000 predicted repellents (0.23% of hits) revealed a diverse group of chemicals that retain some structural features of the known repellents (Fig. 4d, e). We computed partition coefficient ($\log P$) values of the 1,000 compounds to exclude those predicted to be lipophilic ($\log P > 4.5$) and therefore more likely to pass through the skin barrier in topical applications³² (Fig. 4e). We also computed predicted vapour pressures of these chemicals, because volatility may be a useful predictor of spatial volume of repellency (Fig. 4e).

Although the *in silico* screen was feasible, a more significant challenge lies in identifying safe and effective DEET substitutes that can be rapidly approved for human use. To identify such compounds, we applied our *in silico* screen to an assembled natural odour library consisting of >3,000 chemicals identified as originating from plants, insects or vertebrate species, and compounds already approved for human use as fragrances, cosmetics or flavours (Supplementary Information). Although many of the top 200 hits share structural features with known repellents from the training set, they also represent structurally diverse chemicals, allowing targeted exploration of previously untested chemical space (Fig. 4f). For example, several anthranilates and pyrazines were identified, even though such compounds were largely missing from the training set.

Ir40a⁺ cells are activated by new repellents

We selected four compounds from the list: methyl *N,N*-dimethyl anthranilate (MDA), ethyl anthranilate (EA), butyl anthranilate (BA) and 2,3-dimethyl-5-isobutyl pyrazine (DIP), of which the first three have a mild grape-like aroma, excellent safety profiles and have been thoroughly tested and approved for human consumption or oral inhalation by the Food and Drug Administration (FDA), World Health Organization and European Food Safety Authority, and have been listed in the 'generally recognized as safe' (GRAS) list by the Flavour and Extract Manufacturer's Association (Fig. 4g and Supplementary Table 2). The fourth, a pyrazine, is an ant trail pheromone³³. The anthranilate and pyrazine classes also contain a large diversity of chemicals found in nature and therefore present attractive repositories of structural substitutes.

For all four chemicals we found robust activation of sacculus ORNs (Fig. 5a, Supplementary Video 2) that innervate the *Ir40a⁺* 'column' glomerulus (Fig. 5b, as shown for BA). They also activated gustatory neurons that project to similar areas of the SOG as DEET (Fig. 5b, as shown for BA). GCaMP3 imaging in *Ir40a⁺* neurons showed robust responses to these chemicals, whereas several other classes of common

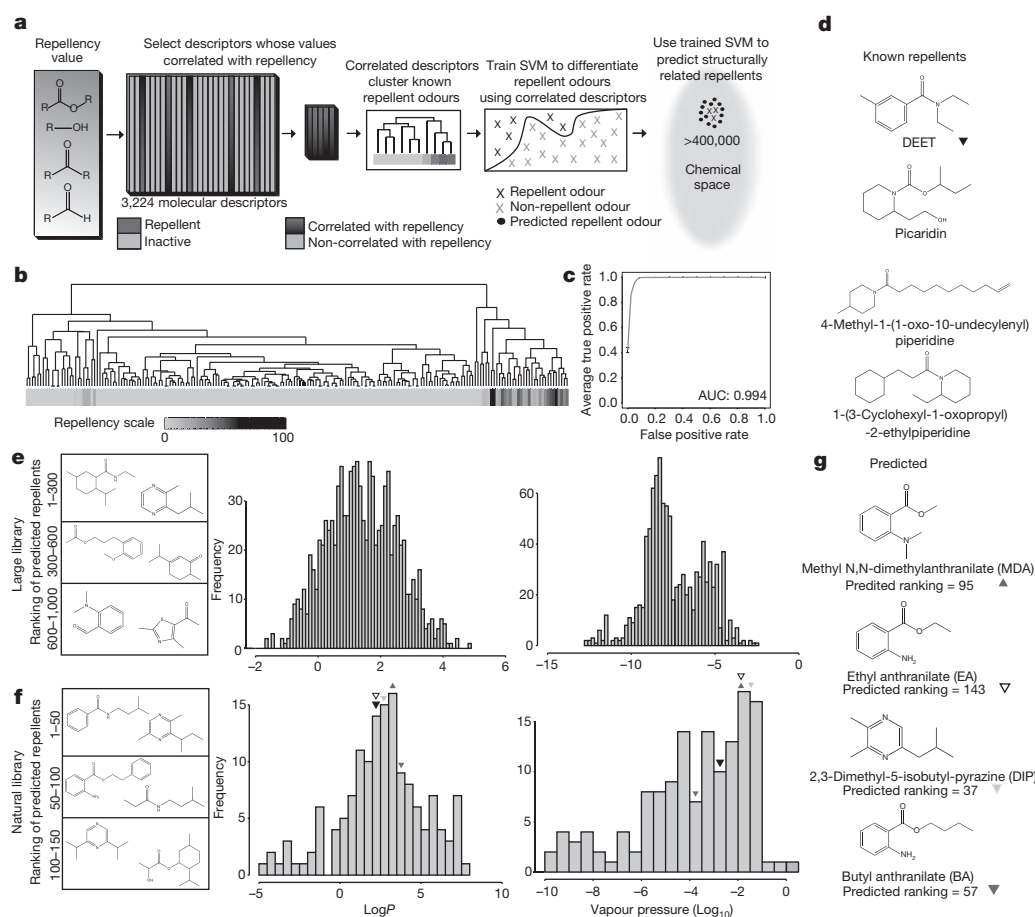


Figure 4 | Chemical informatics prediction of new repellents.

a, Cheminformatics discovery pipeline to identify novel DEET-like repellents. **b**, Hierarchical cluster analysis of 201 training set odorants using optimized descriptors to calculate distances in chemical space. **c**, Receiver operating characteristic (ROC) curve representing computational validation of repellent predictive ability from 20 independent fivefold cross validations. AUC, area under the curve. **d**, DEET, picaridin and two unapproved repellents²⁵. **e**, Representative predicted repellents from >400,000 odorant library (left) and computationally determined values for 1,000 top-ranked predicted repellents (right). **f**, Representative predicted repellents from >3,000 natural odour library (left) and computationally determined values for 150 top-ranked predicted repellents (right). Arrowheads indicate values for DEET and selected odours shown in **g**.

odorants did not (Fig. 5c and Supplementary Fig. 5). These results demonstrate that the computationally predicted chemicals activate the same chemosensory pathways as DEET and are therefore ideal candidates for new repellents.

In order to test the effect of these compounds on behaviour we used a two-choice trap assay in which flies can sense a DEET-treated filter paper positioned at the entrance of a trap through both olfactory and gustatory systems^{3,17} (Fig. 5d). All four compounds had strong dose-dependent repellent effects on *D. melanogaster* (Fig. 5d). Measurements were taken at 24 h and 48 h after the start of the assay, and were found to be consistent. Six additional predicted repellents were tested in a similar manner, at least four of which elicited strong repellency similar to DEET (Supplementary Fig. 6).

To confirm the role of *Ir40a*⁺ neurons in mediating avoidance to these new repellents, we examined behavioural avoidance of flies in which synaptic activity of *Ir40a*⁺ neurons was silenced using TNTG as before. We found that avoidance of chemical treated traps was substantially decreased in *Ir40a*-TNTG flies as compared to control flies (Fig. 5d), showing that *Ir40a*⁺ neurons are required for repellency to the four chemicals.

Mosquitoes avoid predicted repellents

To test the effects of the identified chemicals on mosquito behaviour, we adapted an arm-in-cage assay that allows quantitative analysis of chemical repellency on mosquitoes attracted to a human arm (described in Methods) (Fig. 6a, Supplementary Fig. 7). Female *A. aegypti* mosquitoes showed strong avoidance behaviour to DEET, irrespective of whether or not they could directly contact DEET (Fig. 6b). However, for sporadic landings the average time spent on the net before escape although not significant ($P = 0.203$ for 10% DEET and $P = 0.06$ for 1% DEET, Student's *t*-test) was reduced when direct contact with DEET was permitted, particularly at the lower

concentration (Fig. 6c). Although it is difficult to assess from these experiments the direct contribution of the gustatory system alone, it demonstrates that mosquitoes can avoid DEET strongly at close range, even without making direct contact with it.

In order to test whether the four newly identified *Drosophila* repellents were also olfactory repellents to mosquitoes, we performed behaviour trials using the non-contact version of the assay. Notably, we found that all four compounds applied at 10% concentration demonstrated substantial repellency (Fig. 6d). The fraction of mosquitoes present on the net throughout the duration of the assay (Fig. 6d), as well as the cumulative number of mosquitoes present on the net were substantially decreased in the presence of the test compounds (Fig. 6e). For the mosquitoes that did land on the repellent treatment, the escape index, as measured by the frequency of take-off, was substantially higher as compared to those landing on controls (Supplementary Figs 8 and 9).

One of the major disadvantages of DEET is its property of solubilizing plastics and synthetic materials¹, which affects its usefulness. We tested the ability of the four repellents to dissolve a 3×3 mm square of vinyl. While the vinyl completely disappeared in DEET within 6 h, there was no significant difference in the weight of the vinyl squares immersed in the four DEET substitutes after 6 h or 30 h (Fig. 6f).

Discussion

The unbiased strategy to use a genetic-reporter of neural activity was instrumental in identifying DEET-sensitive *Ir40a*⁺ neurons. These reside in the pit-like sacculus that could protect neurons from harsh chemicals. Both olfactory and gustatory systems are activated by DEET, with additional modes of detection in the antenna being mediated by *orco*¹¹ and a yet to be identified tuning *Or* gene (Fig. 6h). Additionally, DEET has been reported to have a mild enhancing or suppressing effect on the activity of various *Or*-expressing

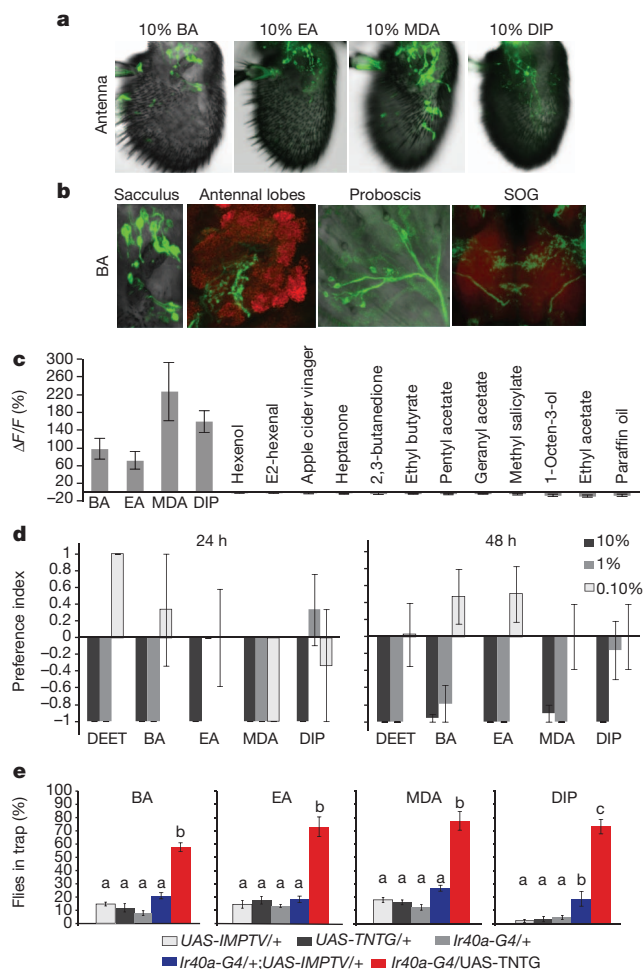


Figure 5 | Predicted repellents activate Ir40a neurons and are strong repellents for *Drosophila*. **a**, Images of antenna of *elav-Gal4/LexAop-CD8-GFP-2A-CD8-GFP*; *UAS-mLexA-VP16-NFAT*, *LexAop-CD2-GFP/+* flies exposed to indicated stimuli for 24 h. **b**, BA-activated GFP⁺ neurons in indicated tissues. **c**, Mean changes in fluorescence intensity in *Ir40a-Gal4/+;UAS-GCaMP3/+* cells after ~2-s application of indicated odors. $n = 9-17$. **d**, Mean responses of flies to predicted repellents in two-choice olfactory and gustatory trap assays measured at 24 h and 48 h. $n = 3-10$ trials (24 h) and 7–10 (48 h); 10 flies per trial, trials with <40% participation were excluded. **e**, Quantification of flies of indicated genotypes entering repellent-treated traps. $n = 6$ trials for each genotype, ~20 flies for each trial. $P < 0.001$, one-way ANOVA with Tukey's post hoc test. For **c–e**, error bars represent s.e.m.

neurons of antennal basicos in *Drosophila*, although a causal relationship between this effect and repellency has not been established⁹. DEET also has a solvent effect that slows down volatile odour release, potentially also from skin¹⁰. Thus, several pathways and mechanisms are likely to participate in overall repellency.

Ir40a can account for the widespread effect of DEET olfactory repellency because it is highly conserved in species that show strong avoidance to it including *Drosophila*, mosquitoes, head lice³⁴ and tribolium³⁵, but not in the honey bee³⁶. *Ir40a* orthologues are conserved across many insects, with several regions of amino acid similarity across the length of the protein (Supplementary Fig. 10). This degree of conservation may better explain the repellent effects of DEET across several insect species compared to *Or* pathways that are not as well conserved. The *Ir40a* pathway therefore has important implications in the development of safe and affordable strategies to control several types of insects and arthropods that are disease vectors of animals and plants or are plant pests.

The chemical informatics enabled us to identify a number of affordable and safe potential repellents that are good candidates for

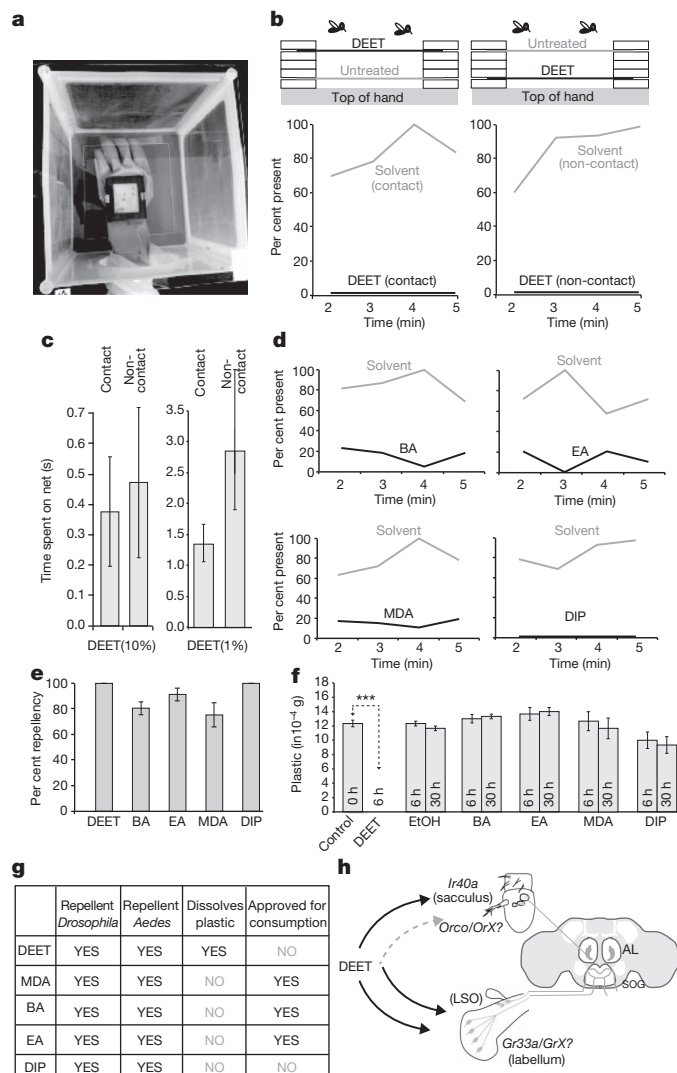


Figure 6 | A new class of mosquito repellents with desirable safety profiles. **a**, Arm-in-cage assay to measure repellency in mosquitoes. **b**, Mean percentage of female *A. aegypti* present for >5 s on top net at indicated times to 10% DEET (black line) or solvent controls performed separately (grey line) in a contact (left) or non-contact (right) assay. **c**, Average time on net for each landing event in **b**. **d**, Mean percentage of female *A. aegypti* present for >5 s on top net in non-contact assay at indicated times. **e**, Cumulative repellency summed across minutes 2–5 of indicated non-contact treatment (10%) in comparison to appropriate solvent control. Forty mosquitoes were used per trial, $n = 5$ trials per treatment for **b–e**. **f**, Mean weight of vinyl pieces following submersion in indicated compounds or ethanol (control) for indicated amount of time. $n = 3$, *** $P < 10^{-5}$, Student's *t*-test. Error bars represent s.e.m. **g**, Properties of new repellents. **h**, Model for DEET detection and processing in *Drosophila*.

regulatory approval for human use (Fig. 6g). This screen identified ~1,000 compounds and >100 additional natural compounds, many approved for use in human food and cosmetics, which may lead to other effective repellents. The repellency strategy may also have promise for use in combination with other behaviour control strategies, such as masking of CO₂-mediated attraction behaviour or population control by trapping as a part of an integrated pull-mask-push strategy^{37,38}. Moreover, these DEET substitutes may be of value in controlling DEET-resistant strains as well. Because several of the new repellents are affordable, activate both the olfactory and bitter gustatory neurons, are approved for human consumption and are strong repellents for fruit flies, they may also have important implications for control of agricultural pest insects that cause enormous crop loss. Novel repellents that are safe and affordable can be used

to limit insect-human contact in disease-endemic areas of the world and to provide an important line of defence against deadly vector-borne diseases.

METHODS SUMMARY

Physiological experiments. NFAT-based neural tracing¹⁶ and GCaMP3-based calcium imaging^{22,23} were performed as previously described with some modifications (see Methods). Single-unit recordings from olfactory sensilla were performed as described previously³⁷.

Behavioural experiments. For olfactory trap assays, 20 *Drosophila* were released in cylindrical arenas containing Eppendorf tube traps (Figs 2d and 3a) with 10% apple cider vinegar as a lure. Repellents were presented on filter papers placed near the trap openings in a manner that did not allow physical contact with the fly before its entering the trap. Trap assays to measure repellency when both olfactory and gustatory inputs were possible were performed as described previously³. Mosquito arm-in-cage avoidance assays were performed with 40 mated *A. aegypti* females held in a cage and presented a human arm that was inserted in a glove containing a window covered with a double-layer of netting. Test compounds were applied to the nettings. Attraction towards the arm was measured using video recordings and analysts were blind to treatments.

Chemical informatics. Optimized molecular descriptors were selected from 3,224 Dragon descriptors based on their ability to increase the correlation between descriptor values and repellency. The repellency-optimized descriptor set was used to first train a support vector machine to predict repellents and then applied to predict new repellents from large compound libraries.

Insects. Fly lines were obtained from the Bloomington *Drosophila* Stock Center for TNT and GCaMP3 experiments, the Vienna *Drosophila* RNAi Center for *UAS-Ir40a* RNAi, J. Wang (UC San Diego) for NFAT tracing, and R. Benton (University of Lausanne) for *Ir40a*-Gal4. Flies were grown on standard cornmeal-dextrose media, at 25 °C unless otherwise noted and mosquitoes at 27 °C and 70% RH.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 January; accepted 15 August 2013.

Published online 2 October 2013.

- Krajick, K. Medical entomology—Keeping the bugs at bay. *Science* **313**, 36–38 (2006).
- Corbel, V. *et al.* Evidence for inhibition of cholinesterases in insect and mammalian nervous systems by the insect repellent DEET. *BMC Biol.* **7**, 47 (2009).
- Reeder, N. L., Ganz, P. J., Carlson, J. R. & Saunders, C. W. Isolation of a DEET-insensitive mutant of *Drosophila melanogaster* (Diptera: Drosophilidae). *J. Econ. Entomol.* **94**, 1584–1588 (2001).
- Klun, J. A. *et al.* Comparative resistance of *Anopheles albimanus* and *Aedes aegypti* to *N,N*-diethyl-3-methylbenzamide (Deet) and 2-methylpiperidinyl-3-cyclohexen-1-carboxamide (AI3–37220) in laboratory human-volunteer repellent assays. *J. Med. Entomol.* **41**, 418–422 (2004).
- Stanczyk, N. M., Brookfield, J. F., Ignell, R., Logan, J. G. & Field, L. M. Behavioral insensitivity to DEET in *Aedes aegypti* is a genetically determined trait residing in changes in sensillum function. *Proc. Natl Acad. Sci. USA* **107**, 8575–8580 (2010).
- Gupta, R. K. & Bhattacharjee, A. K. In *Insect Repellents: Principles, Methods, and Uses* (eds M. Debboun, Frances, S. P. & Strickman, D.) 195–228 (Taylor & Francis Group, 2007).
- Ditzen, M., Pellegrino, M. & Vosshall, L. B. Insect odorant receptors are molecular targets of the insect repellent DEET. *Science* **319**, 1838–1842 (2008).
- Bohbot, J. D. & Dickens, J. C. Odorant receptor modulation: ternary paradigm for mode of action of insect repellents. *Neuropharmacology* **62**, 2086–2095 (2012).
- Pellegrino, M., Steinbach, N., Stensmyr, M. C., Hansson, B. S. & Vosshall, L. B. A natural polymorphism alters odour and DEET sensitivity in an insect odorant receptor. *Nature* **478**, 511–514 (2011).
- Syed, Z. & Leal, W. S. Mosquitoes smell and avoid the insect repellent DEET. *Proc. Natl Acad. Sci. USA* **105**, 13598–13603 (2008).
- DeGennaro, M. *et al.* *orco* mutant mosquitoes lose strong preference for humans and are not repelled by volatile DEET. *Nature* **498**, 487–491 (2013).
- Xia, Y. *et al.* The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl Acad. Sci. USA* **105**, 6433–6438 (2008).
- Liu, C. *et al.* Distinct olfactory signaling mechanisms in the malaria vector mosquito *Anopheles gambiae*. *PLoS Biol.* **8** (2010).
- Lee, Y., Kim, S. H. & Montell, C. Avoiding DEET through insect gustatory receptors. *Neuron* **67**, 555–561 (2010).
- Weiss, L. A., Dahanukar, A., Kwon, J. Y., Banerjee, D. & Carlson, J. R. The molecular and cellular basis of bitter taste in *Drosophila*. *Neuron* **69**, 258–272 (2011).

- Masuyama, K., Zhang, Y., Rao, Y. & Wang, J. W. Mapping neural circuits with activity-dependent nuclear import of a transcription factor. *J. Neurogenet.* **26**, 89–102 (2012).
- Syed, Z., Pelletier, J., Flounders, E., Chitolina, R. F. & Leal, W. S. Generic insect repellent detector from the fruit fly *Drosophila melanogaster*. *PLoS ONE* **6**, e17705 (2011).
- Silbering, A. F. *et al.* Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neurosci.* **31**, 13357–13375 (2011).
- Ai, M. *et al.* Acid sensing by the *Drosophila* olfactory system. *Nature* **468**, 691–695 (2010).
- Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136**, 149–162 (2009).
- Abuin, L. *et al.* Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* **69**, 44–60 (2011).
- Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
- Pelz, D., Roeske, T., Syed, Z., de Bruyne, M. & Galizia, C. G. The molecular receptive range of an olfactory receptor *in vivo* (*Drosophila melanogaster* Or22a). *J. Neurobiol.* **66**, 1544–1563 (2006).
- Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O’Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14**, 341–351 (1995).
- Katritzky, A. R. *et al.* Synthesis and bioassay of improved mosquito repellents predicted from chemical structure. *Proc. Natl Acad. Sci. USA* **105**, 7359–7364 (2008).
- Klocke, J. A., Darlington, M. V. & Balandrin, M. F. 1,8-Cineole (Eucalyptol), a mosquito feeding and ovipositional repellent from volatile oil of *Hemizonia fitchii* (Asteraceae). *J. Chem. Ecol.* **13**, 2131–2141 (1987).
- Kline, D. L., Bernier, U. R., Posey, K. H. & Barnard, D. R. Olfactometric evaluation of spatial repellents for *Aedes aegypti*. *J. Med. Entomol.* **40**, 463–467 (2003).
- Carey, A. F., Wang, G., Su, C. Y., Zwiebel, L. J. & Carlson, J. R. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464**, 66–71 (2010).
- Hallem, E. A. & Carlson, J. R. Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
- Haddad, R. *et al.* A metric for odorant comparison. *Nature Methods* **5**, 425–429 (2008).
- Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
- Walker, J. D., Rodford, R. & Patlewicz, G. Quantitative structure–activity relationships for predicting percutaneous absorption rates. *Environ. Toxicol. Chem.* **22**, 1870–1884 (2003).
- Tentschert, J., Bestmann, H. J., Holldobler, B. & Heinze, J. 2,3-dimethyl-5-(2-methylpropyl)pyrazine, a trail pheromone component of *Eutetratorium mocquersyi* Emery (1899) (Hymenoptera: Formicidae). *Naturwissenschaften* **87**, 377–380 (2000).
- Mumcuoglu, K. Y., Galun, R., Bach, U., Miller, J. & Magdassi, S. Repellency of essential oils and their components to the human body louse, *Pediculus humanus humanus*. *Entomol. Exp. Appl.* **78**, 309–314 (1996).
- Hou, X. W., Fields, P. & Taylor, W. The effect of repellents on penetration into packaging by stored-product insects. *J. Stored Prod. Res.* **40**, 47–54 (2004).
- Abramson, C. I. *et al.* Proboscis conditioning experiments with honeybees, *Apis mellifera caucasica*, with butyric acid and DEET mixture as conditioned and unconditioned stimuli. *J. Insect Sci.* **10**, 1–17 (2010).
- Turner, S. L. & Ray, A. Modification of CO₂ avoidance behaviour in *Drosophila* by inhibitory odorants. *Nature* **461**, 277–281 (2009).
- Turner, S. L. *et al.* Ultra-prolonged activation of CO₂-sensing neurons disorients mosquitoes. *Nature* **474**, 87–91 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Ganguly and D. Carter for help with calcium imaging; Z. Wisotsky for help with gustatory experiments; D. MacWilliam for help with olfactory experiments; J. Wang for sharing the NFAT transgenic fly line; and R. Benton for sharing the *Ir40a-Gal4* fly line. This work was partly funded by a Whitehall Foundation grant to A.D., an R21NS074332 (NINDS) to A.D. and A.R., and an R56AI099778 (NIAID) and R01AI087785 (NIAID) to A.R. The granting agencies had no role in experimental design or analysis.

Author Contributions S.M.B. planned and performed the chemical informatics and solubility experiment, and helped design the behaviour experiments. P.K. planned and performed the NFAT imaging, Ca²⁺ imaging and *Drosophila* behaviour experiments. S.K.T. performed Ca²⁺ imaging, electrophysiology and some behaviour experiments. T.G. performed the arm-in-cage experiments. C.P. performed behaviour analysis. S.M.B., P.K. and S.K.T. helped prepare drafts of the manuscript and figures. A.D. planned experiments and helped write the manuscript. A.R. planned experiments, managed the project, and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.R. (anand.ray@ucr.edu).

METHODS

Fly stocks. Wild type flies were w^{1118} backcrossed to Canton-S for 5 generations. *UAS-GCaMP3* (BL#32236), *UAS-TNTG* (BL#28838), *UAS-IMPTV* (BL#28840) and *Tub-PGal80^{ts}* (BL#7017) were obtained from the Bloomington *Drosophila* Stock Center. The following stocks were generously provided: *LexAop-CD8-GFP-2A-CD8-GFP*; *UAS-mLexA-VP16-NFAT*; *LexAop-CD2-GFP* by J. Wang (UC San Diego, CA), *Ir40a-Gal4* by R. Benton (University of Lausanne, Switzerland), and *elav-Gal4* by L. Luo (Stanford, CA). *UAS-Ir40a* RNAi (line 1) (v101725) and *UAS-Ir40a* RNAi (line 2) (v3960) lines were obtained from the Vienna *Drosophila* RNAi Center. *Ir40a* RNAi is predicted to have no off-targets. Fly stocks were grown on standard cornmeal-dextrose media at 25 °C unless otherwise noted. Flies of appropriate genotypes for behaviour experiments were randomly sorted from populations before performing behavioural or electrophysiological experiments.

NFAT-based neural tracing. Late dark *Drosophila* pupae ready to emerge of genotype *elav-Gal4/ LexAop-CD8-GFP-2A-CD8-GFP*; *UAS-mLexA-VP16-NFAT*; *LexAop-CD2-GFP/+*¹⁶ were collected on moist filter paper strips in culture vials which contained 2 Kimwipes soaked in 5 ml of water in a relatively odour-free environment. A 100 µl sample of odour at the indicated concentration was dissolved in acetone, spread on a filter strip (~1 cm × 3 cm), dried for 1 min and placed in a vial with 10–15 pupae. The exposure was given for 24 h and the filter paper strip with odour was replaced at ~12–14 h with fresh odour.

Calcium imaging using GCaMP3. DEET, DMSO, hexane and candidate compounds were purchased from Sigma-Aldrich or the eMolecules database (<http://www.emolecules.com>) from Enamine, Vitas M Labs or Chembridge and were of the highest purity available. Approximately 10–12-day-old flies raised at 29 °C (to improve Gal4 activity) were anaesthetized and secured by their wings on double-sided sticky tape (ventral side up) on a Petri dish (BD Falcon, 50 × 9 mm). The fly proboscis, head and body were immobilized by sticky tape as shown (Supplementary Fig. 11). One antenna was stably held down using a glass electrode on thin layer of 70% glycerol that enhanced imaging of fluorescence. The antenna was orientated with the arista and sacculus pointing upwards accessible to odours. Odorants were delivered using 5 ml plastic syringes containing 2 Whatman filter paper strips (2 × 3 cm). A fine mist of DEET at indicated concentrations in DMSO was sprayed into the syringe using an atomizer. Fresh atomized odour syringes were prepared immediately before odour delivery. For DEET substitutes (BA, EA, MDA and DIP), a 100 µl of 50% dilution in DMSO was applied to the filter paper directly and for other odorants 100 µl of 10^{−2} solution in paraffin or water for apple cider vinegar (ACV) was applied directly on the filter paper. The odour puff (~2 s) was delivered using the syringe over the antenna manually. For imaging odour-evoked activity from the antenna using GCaMP3, a Leica SP5 inverted confocal microscope was used. A filter block with 488 nm excitation filter and 500–535 nm emission filter was used and images were acquired at 3.3 frames per second with a resolution of 330 × 330 pixels using a 10× objective. The settings were optimized to capture odour-induced responses of GCaMP3 with high spatial and temporal resolution while limiting reporter bleaching.

Data analysis for calcium imaging was performed using the Leica SP5 LAS AF software (in quantify mode) to obtain the heat map images and fluorescence intensity changes. The $\Delta F/F$ percentage was calculated separately for each selected cell body by taking the mean intensity value of all frames for 5 s before the odour puff (F_{pre}) and taking the mean intensity value of all frames for 5 s around the peak responses (F_{post}) after the end of the ~2 s of stimulus delivery period. Similarly, the mean intensity values were taken for a background area in the vicinity of the cells.

The $\Delta F/F$ percentage was calculated according to the formula:

$$\Delta F/F(\%) = \frac{(F_{post} - F_{background(post)}) - (F_{pre} - F_{background(pre)})}{F_{pre} - F_{background(pre)}} \times 100$$

Immunohistochemistry. After 24 h exposure to either odour or solvent (control), flies were anaesthetized on ice and the tissue dissected in chilled 1× PBS and fixed for 30 min in 4% PFA (0.3% Triton X-100) at room temperature. After washes with PBST (PBS with 0.3% Triton X-100) brains were blocked using PBST with 5% bovine serum albumin (BSA). Rabbit anti-GFP (1:1,000, Invitrogen) and anti-nc82 (1:10 Developmental Studies Hybridoma Bank) were used as primary antibodies and samples were incubated for 3 nights at 4 °C. Alexa Fluor 488 anti-rabbit immunoglobulin G (IgG) (Invitrogen; 1:200) and Alexa Fluor 546 anti-mouse IgG (Invitrogen; 1:200) were used as secondary antibodies, respectively, followed by overnight incubation at 4 °C. Images were acquired with a Zeiss or Leica SP5 confocal microscope and images processing was done using ImageJ and Photoshop software. Data analysis was performed offline, and the

investigator was blind to the treatment while counting GFP⁺ antennal neurons in the confocal micrographs.

Temperature sensitive Gal80^{ts} experiment. For the two-choice behaviour assay in Fig. 3 and Supplementary Fig. 4, flies (10 males and 10 females) with genotypes *Ir40a-Gal4/+*; *UAS-Ir40a* RNAi(2)/*Gal80^{ts}* were grown throughout at 18 °C (permissive temperature) where Gal80 is active and RNAi is off. Such flies were treated as control. In parallel, flies of the same genotype were shifted to 29 °C (non-permissive temperature) from 18 °C as late black pupae for 4 days to activate Gal4 and switch on RNAi. These flies were used as knockdown flies. A subset of flies that were shifted to 29 °C was shifted back to 18 °C for 4 additional days to turn off the RNAi and these were used as recovery flies.

Electrophysiology. Flies used were 4–7-days-old and raised on cornmeal food at 25 °C. Extracellular recordings were made by inserting a glass electrode into the base of a palp sensillum as done previously^{37,39}. Odorants were diluted in hexane or DMSO, at indicated concentrations (made fresh for every stimulus). For DEET stimulation, 10 µl of diluted odorant was applied to a filter paper strip, the hexane solvent was evaporated for 30 s (as in a previous study¹⁷) or for 5 min, and placed into a glass pasture pipette cartridge, and each cartridge was only used once. The evaporation of hexane from the filter paper strip was much slower upon mixing with DEET and lingering dampness of the filter paper could be observed visually as well.

Behavioural testing of *Drosophila* olfactory avoidance assay for DEET. For each trial, flies that were 3–6 days old (10 males and 10 females) were starved for 18 h.

For the trap assay, flies were transferred to a cylindrical 38.1 mm (diameter) × 84.1 mm (height) chamber containing a trap fashioned from an upturned 1.5 ml microcentrifuge tube with 2 mm removed from the tapered end. A pipette tip (1,000 µl) was cut 2.5 cm from the narrow end and 0.5 cm from top and inserted into the bottom of the inverted microcentrifuge tube. A 15 mm × 16 mm #1 Whatmann filter paper was inserted in between the pipette tip and tip of microcentrifuge tube so that entering flies could not make physical contact with it. A 25 µl sample of test compound was applied to filter paper and 125 µl of 10% ACV was applied to the upturned lid of the microcentrifuge tube as attractant. Trials were run for 24 h and the numbers of flies entering the trap counted (Fig. 2d).

In the two-choice test, two 10% ACV (125 µl) lured traps as described above were placed in the cylinder, one with 50 µl solvent (DMSO) and another with 50 µl the test odorant at 50% applied to the filter paper (Fig. 3). The more volatile DIP was tested at a lower concentration of 25%. For positive control tests in Supplementary Fig. 4, 125 µl of 10% ACV in test traps and 125 µl of water in control traps was added in the upturned microcentrifuge tube lid. Both traps contained filter papers as before with 50 µl solvent (DMSO). All trials were run for 24 h, positions randomized, and counted. Only trials with >35% participation were considered.

$$\text{Preference index} = \frac{\text{Number of flies in treated trap} - \text{number in control trap}}{\text{Number of flies in treated} + \text{control traps}} \quad (1)$$

***Drosophila* olfactory and gustatory avoidance assay for DEET.** Repellency was tested in Fig. 5d and Supplementary Fig. 6 using a *Drosophila melanogaster* two-choice trap assay as described previously^{3,17} with minor modifications. Briefly, traps were made with two 1.5 ml microcentrifuge tubes (USA Scientific) and 20 ml pipette tips (USA Scientific), each cap contained standard cornmeal medium. A T-shaped piece of filter paper (Whatman #1) was impregnated with 5 µl of acetone (control) or 5 µl of 10%, 1% or 0.10% test odour, diluted in acetone. Traps were placed within a Petri dish (100 × 15 mm, Fisher) containing 10 ml of 1% agarose to provide moisture. Ten wild-type Canton-S flies 4–7-days-old were used per trial, which lasted 48 h, by which time point nearly all flies in the assays had made a choice. For the 24 h time point data were considered only if >35% of flies had made a choice; at 48 h the majority of flies had made choices. The preference index was calculated as in equation (1) above.

Mosquito arm-in-cage avoidance assay for DEET. Repellency was tested in mated and starved *A. aegypti* females using an arm-in-cage assay. *A. aegypti* mosquitoes (eggs obtained from Benzon Research) were maintained at ~27 °C and 70% relative humidity on 14 h:10 h light:dark cycle. Behavioural tests were done with 40 mated, non-blood fed, ~24 h starved, 4–10-day-old females in 30 cm × 30 cm × 30 cm cages with a glass top to allow for video recording (Fig. 6a, Supplementary Fig. 7). The experimental protocol was reviewed and approved by the Institutional Review Board (IRB) Compliance Analyst at UCR and determined not to require additional Human Research Review Board approval. Each test compound solution (500 µl) of 10% concentration in acetone solvent was applied evenly to a white rectangular 7 cm × 6 cm polyester netting (mesh size 26 × 22 holes per square inch) in a glass Petri dish and suspended in the air for 30 min to allow solvent evaporation. The more volatile 2,3-dimethyl-5-isobutyl

pyrazine was dissolved in paraffin oil. Acetone or paraffin oil (500 µl) served as control. A nitrile glove (Sol-vex) was modified as described in Supplementary Fig. 7 such that a 5.8 cm × 5 cm window was present for skin odour exposure. A set of magnetic window frames were designed to secure the treated net ~1.5 mm above skin, and a second untreated netting ~4.5 mm above the treated net in a manner so that mosquitoes were attracted to skin emanations in the open window but unable to contact treated nets with tarsi, or contact and pierce skin. Additionally the test compound had minimal contact with skin. A clean set of glove and magnets were used for every trial. Care was taken that the experimenter did not use cosmetics such as soap on arms. For each trial the arm was first inserted for 5 min and the number of mosquitoes landing or escaping test window recorded on video for a 5-min period. Solvent controls were always tested before a treatment. Mosquitoes showed robust attraction to a solvent treated arm when offered a second time after a gap of 5 mins providing a rigorous test for the treatments to be tested second. No cage was tested more than once within 1 h of a testing session and not more than twice on any single day. Videos were analysed blind and the numbers of mosquitoes present for a 5-s continuous duration were counted every minute. Mosquitoes reliably started accumulating in controls at the 2 min point and data from this time point were considered for analysis.

Percentage present was calculated as the average number of mosquitoes on the window for 5 s at a given time point across trials. All values were normalized to percentage of the highest value for the comparison, which was assigned a 100 per cent present.

Percentage repellency = $(1 - (\text{mean cumulative number of mosquitoes on the window of treatment for 5 s at time points 2, 3, 4, 5 min} / \text{mean cumulative number of mosquitoes that remained on window of solvent treatment for 5 s at time points 2, 3, 4, 5 min})) \times 100$.

Escape index = $(\text{average number of mosquitoes in treatment that landed yet left the mesh during a five second window over the following time points: 2 min, 3 min, 4 min, 5 min} / (\text{average number of mosquitoes that landed yet left the mesh during a 5 s window over the same time points in (treatment + control)}))$

Each time point had $n = 5$ trials, 40 mosquitoes per trial, except for EA, in which $n = 4$.

Chemical Informatics. A single energy-minimized three-dimensional structure was predicted for each compound using the Omega2 software package⁴⁰. The commercially available software package Dragon (3,224 individual descriptors) from Taleté was used to calculate molecular descriptors⁴¹. Descriptor values were normalized across compounds to standard scores by subtracting the mean value for each descriptor type and dividing by the standard deviation. Molecular descriptors that did not show variation across compounds were removed.

For our analysis, compounds from different studies were approximated into a single metric of 'protection duration' as a rough indicator of repellency. The non-repellent diversifying training set of odours were assigned protection times of zero, whereas the approved repellents DEET and picaridin were assigned the highest value since we made the assumption that these would have structural properties important for regulatory approval. Compounds were clustered using Euclidean distance and hierarchical clustering based on differences in repellency values, and a set of 5 compounds with the highest activity that clustered together was classified as 'training repellents'.

A compound-by-compound repellency distance matrix was calculated from repellency data. A separate compound-by-compound descriptor distance matrix was calculated using the 3,224 descriptor values calculated by the Dragon software package. Using a sequential forward selection (SFS) approach, all descriptors are individually compared and selected for their ability to increase the correlation between descriptor values and repellency. The descriptor that correlates best is retained and each further iteration adds an additional descriptor to improve the correlation values. This process is continued until additional descriptors fail to improve the correlation value from the previous step. This process results in a unique descriptor set that is optimized for repellency.

This repellency-optimized descriptor set was used to train a support vector machine (SVM) using regression and a radial basis function kernel available in the R package e1071, which integrates libsvm^{42,43}. Optimal gamma and cost values were determined using the tune.SVM function. The resulting trained SVM was

then applied to predict activity for compounds from two libraries *in silico*, a natural compound library of ~3,200 volatiles and a > 440,000 compounds library.

For the natural compound library we assembled a subset of 3,197 volatile compounds from defined origins including plants, humans, insects⁴⁴, food flavours and a fragrance collection⁴⁵ including fruit and floral volatiles^{46–53}. For the larger library we assembled a subset of >440,000 small molecules from the eMolecules database that have properties of volatile odourants. (Molecular weight <325 grams per mole and atoms: C, O, N, H, S.)

We performed a fivefold cross-validation by dividing the data set randomly into five equal sized partitions. Four of the partitions were applied to train the SVM and the remaining partition, which was not used for training, was used to test predictive ability. This process was repeated five times, each trial excluding a different subset of compounds as the training set and assigning the remainder as the test set. The whole process was repeated 20 times to improve consistency. A receiver operating characteristics (ROC) analysis was then used to analyse the performance of our computational repellency prediction. The overall predictive ability was calculated as a single receiver operating characteristic (ROC) curve for all 20 independent validations.

Calculation of LogP and vapour pressure values. SMILES structures of the predicted repellent odours were used with EPI Suite (<http://www.epa.gov/oppt/exposure/pubs/episuite.htm>) to calculate predicted LogP and vapour pressure values.

Vinyl solubility test. One 3 × 3 mm square of 4 gauge vinyl was submerged in 1 ml of each test compound in a glass container, stirred at a constant rate on a shaker and checked every 30 min until the vinyl square in DEET was completely dissolved (6 h). The vinyl pieces in each of the other compounds were removed, rinsed in ethanol and weighed. The process was repeated at 30 h (24 h after the vinyl square completely disappeared in DEET).

Statistical analyses. For behaviour experiments with preference index, arcsine-transformed data were analysed. Tests used are indicated in the figure legends and they are Student's *t*-test, one-way ANOVA and Tukey's post hoc analysis. Statistical tests for each experimental category and sample trail sizes were selected on the basis of previously published studies using similar assays, which are cited throughout the manuscript. For all graphs, error bars indicate s.e.m.

39. de Bruyne, M., Clyne, P. J. & Carlson, J. R. Odor coding in a model olfactory organ: the *Drosophila* maxillary palp. *J. Neurosci.* **19**, 4520–4532 (1999).
40. Boström, J., Greenwood, J. R. & Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph Model.* **21**, 449–462 (2003).
41. DRAGON software for Windows for molecular descriptor calculations v.5.5 (Taleté, 2007).
42. Chang, C. & Lin, C. Libsvm: a library for support vector machines. (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>) (2001).
43. Karatzoglou, A., Meyer, D. & Hornik, K. Support vector machines in R. *J. Stat. Softw.* <http://www.jstatsoft.org/v15/i09> (2006).
44. El-Sayed, A. M. The Pherobase database of insect pheromones and semiochemicals <http://www.pherobase.com/> (2009).
45. Flavors and fragrances 2007–2008 catalogue (Sigma-Aldrich, 2007).
46. Cork, A. & Park, K. C. Identification of electrophysiologically-active compounds for the malaria mosquito, *Anopheles gambiae*, in human sweat extracts. *Med. Vet. Entomol.* **10**, 269–276 (1996).
47. Curran, A. M., Rabin, S. I., Prada, P. A. & Furton, K. G. Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. *J. Chem. Ecol.* **31**, 1607–1619 (2005).
48. Gallagher, M. *et al.* Analyses of volatile organic compounds from human skin. *Br. J. Dermatol.* **159**, 780–791 (2008).
49. Knudsen, J. T., Eriksson, R., Gershenzon, J. & Ståhl, B. Diversity and distribution of floral scent. *Bot. Rev.* **72**, 1–120 (2006).
50. Logan, J. G. *et al.* Identification of human-derived volatile chemicals that interfere with attraction of *Aedes aegypti* mosquitoes. *J. Chem. Ecol.* **34**, 308–322 (2008).
51. Meijerink, J. *et al.* Identification of olfactory stimulants for *Anopheles gambiae* from human sweat samples. *J. Chem. Ecol.* **26**, 1367–1382 (2000).
52. Zeng, X. N. *et al.* Analysis of characteristic odors from human male axillae. *J. Chem. Ecol.* **17**, 1469–1492 (1991).
53. Zeng, X. N., Leyden, J. J., Spielman, A. I. & Preti, G. Analysis of characteristic human female axillary odors: qualitative comparison to males. *J. Chem. Ecol.* **22**, 237–257 (1996).

Spatial organization within a niche as a determinant of stem-cell fate

Panteleimon Rombolas¹, Kailin R. Mesa¹ & Valentina Greco¹

Stem-cell niches in mammalian tissues are often heterogeneous and compartmentalized; however, whether distinct niche locations determine different stem-cell fates remains unclear. To test this hypothesis, here we use the mouse hair follicle niche and combine intravital microscopy with genetic lineage tracing to re-visit the same stem-cell lineages, from their exact place of origin, throughout regeneration in live mice. Using this method, we show directly that the position of a stem cell within the hair follicle niche can predict whether it is likely to remain uncommitted, generate precursors or commit to a differentiated fate. Furthermore, using laser ablation we demonstrate that hair follicle stem cells are dispensable for regeneration, and that epithelial cells, which do not normally participate in hair growth, re-populate the lost stem-cell compartment and sustain hair regeneration. This study provides a general model for niche-induced fate determination in adult tissues.

Stem-cell niches in adult tissues constitute a spatially distinct micro-environment, including neighbouring cells, signals and extracellular material^{1,2}. Anatomical and molecular heterogeneity seems to be a common feature between mammalian stem-cell niches across different tissues^{3–6}; however, it is unclear whether the specific location that a stem cell occupies within the niche can influence its function.

Haematopoietic stem cells with divergent roles in homeostasis and pathophysiology are proposed to associate with distinct niche compartments, such as the endosteum or the vasculature in the central bone marrow, which may affect their behaviour and possibly their long-term fate^{7–13}. Moreover, in the intestinal crypt, fast-cycling stem cell and quiescent progenitor populations reside in distinct positions at the bottom of the crypt, whereas the transient amplified pool of precursors line the walls of the crypt, progressively differentiating as they reach the surface of the villi^{14–17}. A neutral competition model has been proposed for long-term homeostasis of the intestinal niche^{18,19}, but the short-term behaviour of individual stem cells in different positions at the bottom of the crypt has not been determined. The hair follicle in the skin represents another highly compartmentalized niche where stem cells reside in the bulge, while a pool of progenitors, called the hair germ, is clustered in a different niche location directly below (Extended Data Fig. 1a)^{20–24}. This common theme of niche compartmentalization in the above examples raises the question as to whether stem cells within their compartments are functionally equivalent. Specifically, it is not clear whether each stem cell can stochastically generate every lineage in a tissue, or whether the precise position within the niche can impose a distinct fate.

The mouse hair follicle is a self-contained mini-organ that represents a unique system for monitoring niche behaviour *in vivo*, because the location of stem cells and differentiated cell types is anatomically distinct and molecularly well defined^{20–22,25–31} (Extended Data Fig. 1a, b). Hair follicles normally undergo stereotypic cycles of regeneration, which in the young mouse are highly synchronized across large areas of the skin and therefore the exact timing of rest, growth and regression phases can be accurately predicted³² (Extended Data Fig. 1c). During hair growth, mesenchymal–epithelial crosstalk at the bottom of the hair follicle niche induces the formation and upwards expansion of seven concentric differentiated layers. These inner layers make up the hair shaft and supportive

inner root sheath (IRS), whereas a relatively undifferentiated outer cell layer called the outer root sheath (ORS) grows downwards to envelop the elongating hair follicle fully^{29,31,33–37} (Extended Data Fig. 1b). Taking advantage of the accessibility of the skin hair follicle, we previously established the ability to visualize these processes non-invasively, *in vivo*²⁴. Here, we have developed a new approach to mark single stem cells in different positions within the niche and re-visit the same lineages over a period of several weeks to months, in live mice. Furthermore, we use laser-induced cell ablation to test whether hair follicle stem cells are required for hair regeneration and to address how injury-induced cell mobility between different niches affects their fate.

Niche location predicts stem-cell fate

To explore the significance of specific niche positioning to stem-cell fate, we implemented an *in vivo* lineage tracing approach at single-cell resolution by live imaging (Extended Data Fig. 2). To mark hair follicle stem cells in the bulge and hair germ compartments genetically, we used mice containing either *K19-CreER* (expressing tamoxifen-inducible Cre; also known as *Krt19-CreER*) or *Lgr5-CreER*, in addition to *Rosa-stop-tdTomato* reporter alleles^{16,38–41} (Extended Data Fig. 2a). *K14-H2BGFP* (histone H2B fused with green fluorescent protein (GFP) and driven by the *Krt14* promoter) and *Lef1-RFP* (red fluorescent protein (RFP) driven by the *Lef1* promoter) were used as general epithelial and mesenchymal fluorescent reporters as described previously²⁴ (Extended Data Figs 1a and 3a). Mice were induced with a single low dose of tamoxifen in the first rest phase of the hair cycle (first telogen, approximately postnatal day (P) 20) and stem cells were visualized *in vivo* three days later (~P23), while the hair follicles were still quiescent (Extended Data Fig. 3a). We verified that marked cells did not translocate from their initial position within the niche and that no additional ectopic expression of the Cre reporter occurred owing to Cre recombinase leakage while hair follicles remained quiescent (Extended Data Fig. 3b). As hair regeneration commenced, we re-visited the same follicles in separate imaging sessions and the lineage progression of previously identified single stem cells was documented (Fig. 1).

Analysis of the *in vivo* lineage tracing data showed that during this process the fate of individual stem cells followed highly stereotypic patterns, which correlated with their original location within the niche

¹Department of Genetics, Department of Dermatology, Yale Stem Cell Center, Yale Cancer Center, Yale School of Medicine, New Haven, Connecticut 06510, USA.

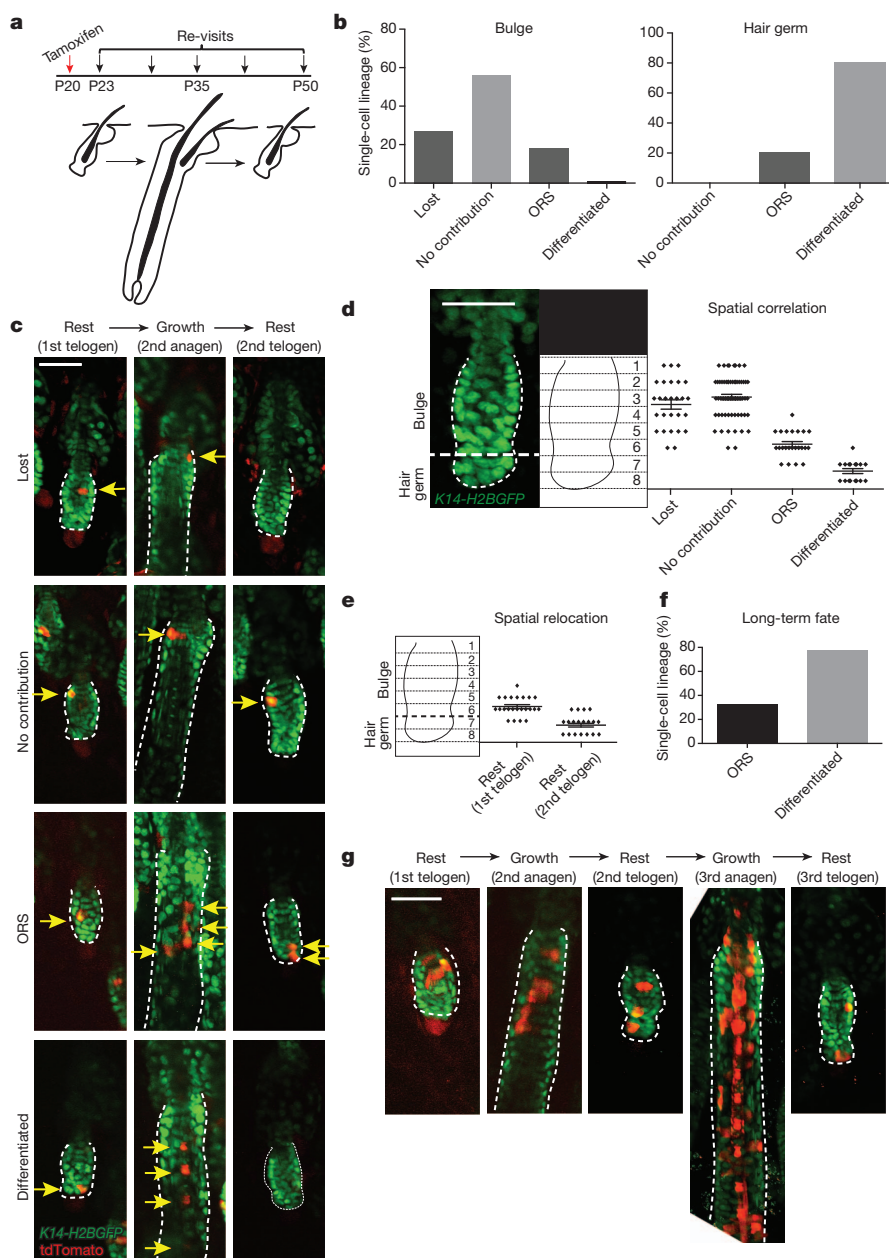


Figure 1 | Niche location can predict the fate of hair follicle stem cells. **a**, Scheme of single stem-cell lineage tracing in live mice. **b**, Statistical analysis of the fate of stem cells originating from the bulge or hair germ ($n = 108$ or 20 lineages, respectively, in 8 mice). **c**, Representative examples of single stem-cell lineages (arrows) traced during a full hair cycle. Each sequence represents a different fate that correlates with a specific niche location. **d**, Graphical correlation between the original location of a single stem cell and its fate after a full hair cycle ($n = 128$; error bars represent s.e.m.). **e**, Spatial relocation of ORS lineages after a full hair cycle ($n = 23$; error bars represent s.e.m.). **f**, Quantification of the fate of ORS lineages in the second hair cycle ($n = 9$). **g**, *In vivo* lineage tracing of bulge stem cells over two consecutive hair cycles. Scale bars, 50 μm .

at the onset of a regeneration cycle (Fig. 1b–d). Specifically, most of the stem cells located within the bulge did not contribute to the subsequent hair cycle or were lost, and a smaller fraction of bulge stem cells produced lineages only in the relatively undifferentiated outer layer (ORS) (Fig. 1b, c, Extended Data Figs 4 and 5, and Supplementary Video 1). Conversely, cells located in the hair germ consistently contributed to hair follicle growth by generating differentiated lineages (Fig. 1b, c, Extended Data Figs 4 and 6, and Supplementary Video 1). Even within each niche compartment the precise location dictated different stem-cell behaviours. For example, within the bulge, stem cells situated in the lower half of the compartment were more likely to proliferate and generate ORS lineages than stem cells situated in the upper half, which were either quiescent or generated limited clones that remained in the bulge (Fig. 1d). These data show a direct correlation between a specific niche location and stem-cell fate.

To test the long-term fate of hair follicle stem cells, we traced bulge lineages over two consecutive hair cycles. Bulge stem cells that persisted in the upper portion of the bulge compartment after the first cycle remained there during the second cycle (Extended Data Fig. 7).

However, lower bulge descendants that acquired an ORS fate were often found in the hair germ after hair follicle regression and entry into the next rest phase (second telogen; Fig. 1e and Extended Data Fig. 8). These ORS clones more frequently gave rise to differentiated lineages in the second hair cycle, consistent with their new position in the hair germ (Fig. 1f). Thus, bulge stem cells can contribute to hair growth by following a stepwise transition to differentiation through an intermediate ORS fate (Fig. 1g). These data also enforce the notion that fate is established at the onset of a new regeneration cycle depending on the specific location that a stem cell occupies in the niche.

ORS expansion is spatially regulated

Our data suggest that the ORS represents an intermediate stage between quiescent bulge stem cells and hair germ cells. Notably, lineage tracing indicated that ORS clones often expanded discontinuously towards the bulb (Fig. 2a, Extended Data Fig. 8 and Supplementary Video 2). To understand how the niche influences this mode of ORS expansion, we collected several time-lapse recordings of hair follicles in advanced growth stages (anagen III–IV; Supplementary Videos 3 and 4). Analysis of cell

behaviour at this stage of growth showed that the ORS undergoes a spatially regulated mode of expansion. Specifically, cell proliferation was restricted to a narrow zone between the lower bulge and the bulb (Fig. 2b and Supplementary Videos 3 and 4). Cell divisions in this 'proliferative zone' were highly oriented, with a mitotic spindle perpendicular to the long axis of growth (Fig. 2b). These oriented cell divisions may not contribute directly to the longitudinal expansion of the ORS, and as a result this proliferative zone displayed higher cell density than other areas (Fig. 2c, d and Supplementary Videos 3 and 4). Further analysis of the time-lapse videos revealed that cells at the distal border of the proliferative zone became mobile and migrated rapidly towards the bulb, thus directly contributing to the downward expansion of the ORS. This previously uncharacterized mode of cell migration indicates highly dynamic cell–cell contacts and may partially explain the discontinuous appearance of the ORS clones observed by lineage tracing (Fig. 2e, f and Supplementary Video 5). This bimodal type of ORS growth and the spatially defined areas of proliferation and migration highlight the regional control that the niche exerts during growth.

Bulge stem cells are dispensable

Our lineage tracing experiments suggest a functional compartmentalization of the niche, in which stem cells positioned in the lower bulge may specify the ORS and those in the hair germ the differentiated hair lineages. To test the stringency of niche-imposed fates towards hair regeneration, we used laser-induced cell ablation to remove specifically either the bulge or the hair germ at the onset of hair growth (first telogen, ~P20; Fig. 3a). To recognize each targeted compartment we used reliable anatomical features of the niche²⁴ (Extended Data Fig. 1a), because available genetic markers label overlapping populations that extend across both the bulge and hair germ^{28,40,41}. Notably, after ablation of either the bulge or the hair germ, the niche consistently recovered the lost cell population, regained its anatomical features and proceeded with hair regeneration (Fig. 3b, c). To verify the efficiency of the ablation process we re-visited the ablated follicles shortly after ablation. Instances in which bulge or hair germ ablation impaired hair regeneration were the result of extensive damage that affected the entire niche and/or the mesenchyme (dermal papilla), consistent with previous reports^{24,29,42} (Fig. 3c). However, some such examples provided crucial information on the dynamics between the epithelium and the mesenchyme. For

instance, when the epithelium and the mesenchyme were physically separated as the result of laser ablation, both the full recovery of the niche and hair regeneration was impaired, even though the mesenchymal dermal papilla lingered a few micrometers below (Fig. 3d). Conversely, in some follicles where the bulge was ablated the hair germ was able to initiate growth, encompassing the dermal papilla before the full recovery of the bulge (Fig. 3e). Overall, our data show that the bulge and hair germ populations are mutually dispensable for hair regeneration as long as a functional interaction between the epithelium and the mesenchymal dermal papilla is maintained (Fig. 3f). Furthermore, they suggest that the ability of the hair germ to initiate hair growth may occur independently of bulge input.

Cell fate changes on niche injury

To explore the cellular mechanisms of niche recovery, we performed time-lapse recordings shortly after bulge laser ablation. The hair germ became proliferative consistent with previous experiments that show hair germ contribution to the niche after stem-cell depletion due to plucking²³. Notably, distant epithelial cells above the bulge (infundibulum) were also observed to become proliferative, and some cells descended rapidly into the niche (Supplementary Videos 6 and 7). These findings raised the possibility that neighbouring epithelial cells situated above the bulge may contribute to the recovery of the niche. To test this hypothesis, we implemented our *in vivo* lineage tracing approach to monitor the behaviour of cells outside the hair follicle niche after bulge ablation. To mark the outermost epithelial layers located above the bulge exclusively we took advantage of the particular expression profile of *K14-CreER/Rosa-stop-tdTomato* mice, in which labelling is strongly biased towards the interfollicular epidermis, infundibulum and sebaceous glands (Fig. 4a and Supplementary Video 8). After induction, follicles that did not contain any labelled cells within the niche were targeted for bulge ablation (Fig. 4a, b). In the days after the ablation there was a significant influx of labelled epithelial cells into the niche, in contrast to neighbouring non-ablated follicles where no additional tdTomato⁺ cells appeared to enter the hair follicle (Fig. 4b).

We found that these 'new' niche cells not only contributed to re-establishing the lost bulge compartment but also participated in the subsequent hair growth, suggesting that they acquired a different fate on assuming their new position in the hair follicle niche (Fig. 4b). To

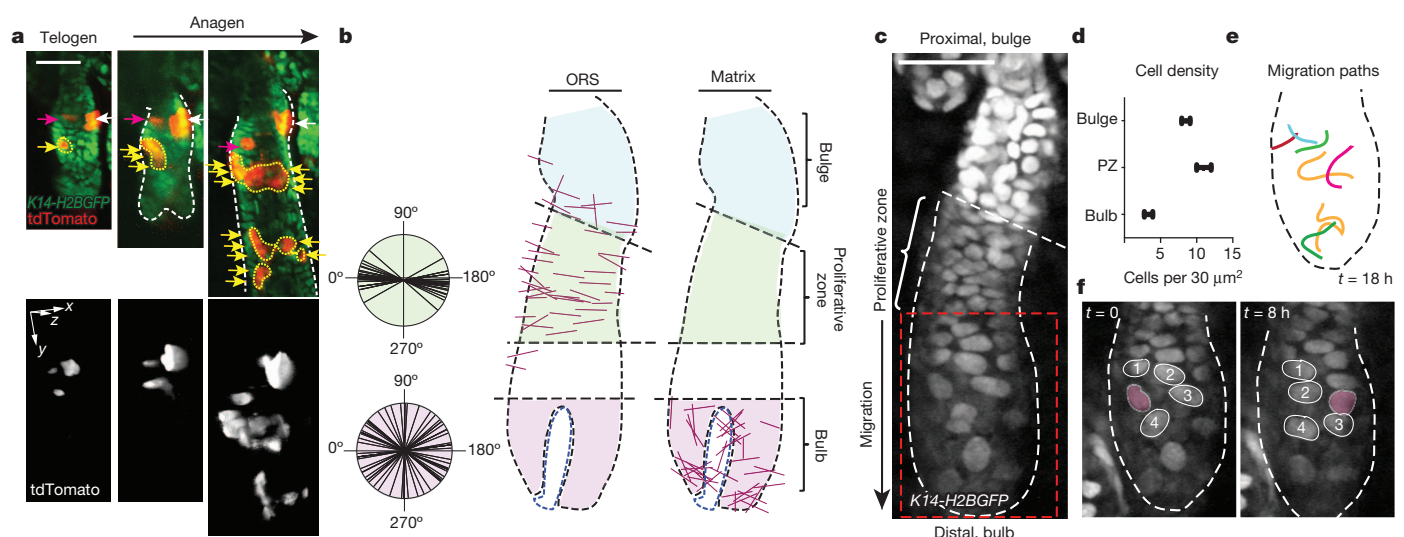


Figure 2 | Mode of ORS growth. **a**, *In vivo* lineage tracing sequence (top) and corresponding three-dimensional renderings of the Cre reporter (bottom) showing ORS expansion during hair growth. Arrows denote cell lineages in different colours. **b**, Graphical representation of the location and axis of cell divisions in the ORS and matrix (bulb) in advanced hair follicle growth (anagen III–IV). **c**, ORS cell distribution during active hair growth. **d**, Quantification of

cell density in different regions of the outer hair follicle epithelial layer (n = 9 per region; error bars represent s.e.m.). **e**, Individual traces of migrating ORS nuclei (see also Supplementary Video 3). **f**, Traces of ORS nuclei depicting their relative positions at two time-points 8 h apart (see also Supplementary Video 3). Scale bars, 50 μm.

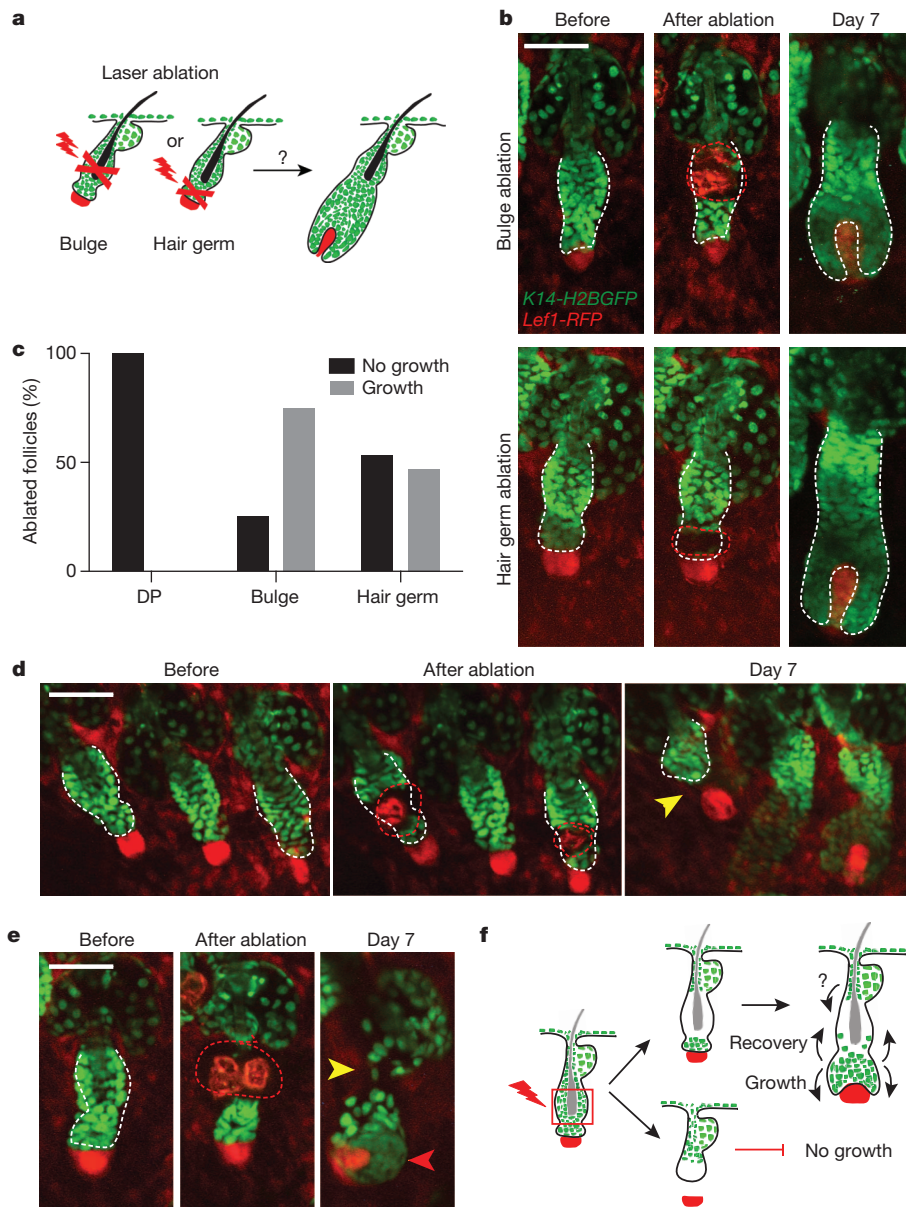


Figure 3 | Bulge and hair germ are mutually dispensable for hair regeneration. **a**, Scheme of the laser ablation experiment. **b**, Sequential snapshots of hair follicles immediately or a week after bulge and hair germ ablation. Red dashed line represents the ablated region. **c**, Quantification of the regenerative capacity of follicles with ablated dermal papilla (DP), bulge and hair germ ($n = 30, 32$ and 28 follicles, respectively, in 9 mice). **d**, Example of hair growth impairment due to the physical separation of the epithelium from the mesenchymal dermal papilla after ablation. **e**, Example of hair follicle growth initiation before full niche recovery. **f**, Scheme of hair follicle niche responses after laser ablation. Scale bars, 50 μm .

test further whether the niche influences the same type of behaviour on the epithelial cells that re-populated the bulge, we used label retention to analyse for quiescence, a hallmark of bulge stem cells^{20,21}, during the second growth cycle after bulge ablation (Fig. 4c). At full growth (third anagen) the bulge of ablated hair follicles displayed significant label retention compared to the lower growing portion of the follicle, but similar to the bulge of non-ablated neighbouring follicles (Fig. 4d, e). Thus, these data provide direct evidence that loss of a stem cell pool due to injury can induce neighbouring epithelial cell populations that do not normally have a hair follicle fate to be mobilized and contribute to re-establishing the niche anatomically as well as functionally. Most importantly, once these cells enter the niche they display characteristics consistent with a hair follicle fate enacted on them in their new location.

Discussion

The relationship between niche position and stem-cell fate is a fundamental question in mammalian stem-cell biology that has remained unanswered. Current approaches to address this problem involve the use of genetic lineage tracing tools based on inducible Cre recombinase,

driven by stem-cell-specific promoters⁴³. However, the mosaic expression of the Cre reporter within the stem-cell pool and the inability to follow individual stem cells over time has greatly limited our understanding of the fate of individual cells at precise locations within the niche. To overcome these limitations we have devised a system that combines genetic lineage tracing with intravital microscopy to monitor the progression of single stem-cell lineages from their initial position, by re-visiting the same undisturbed niche in separate experiments in live mice.

Using this approach, we found evidence that establishes a strong link between a specific niche location and stem-cell fate. Although a cell-autonomous model is plausible, our data support a model for fate determination in the hair follicle that is based on the spatial organization of the niche (Extended Data Fig. 9). According to this model, a cell in the upper half of the bulge is favoured to remain uncommitted to a specific fate and therefore more likely to remain quiescent or self-renew. By contrast, a cell situated in the lower bulge will be subject to activating stimuli from the niche driving it to undergo limited amplification as part of the still relatively undifferentiated ORS. The fraction of the ORS pool that survives the regression phase of the hair cycle will now be situated in the compartment that becomes the new hair germ. Once in

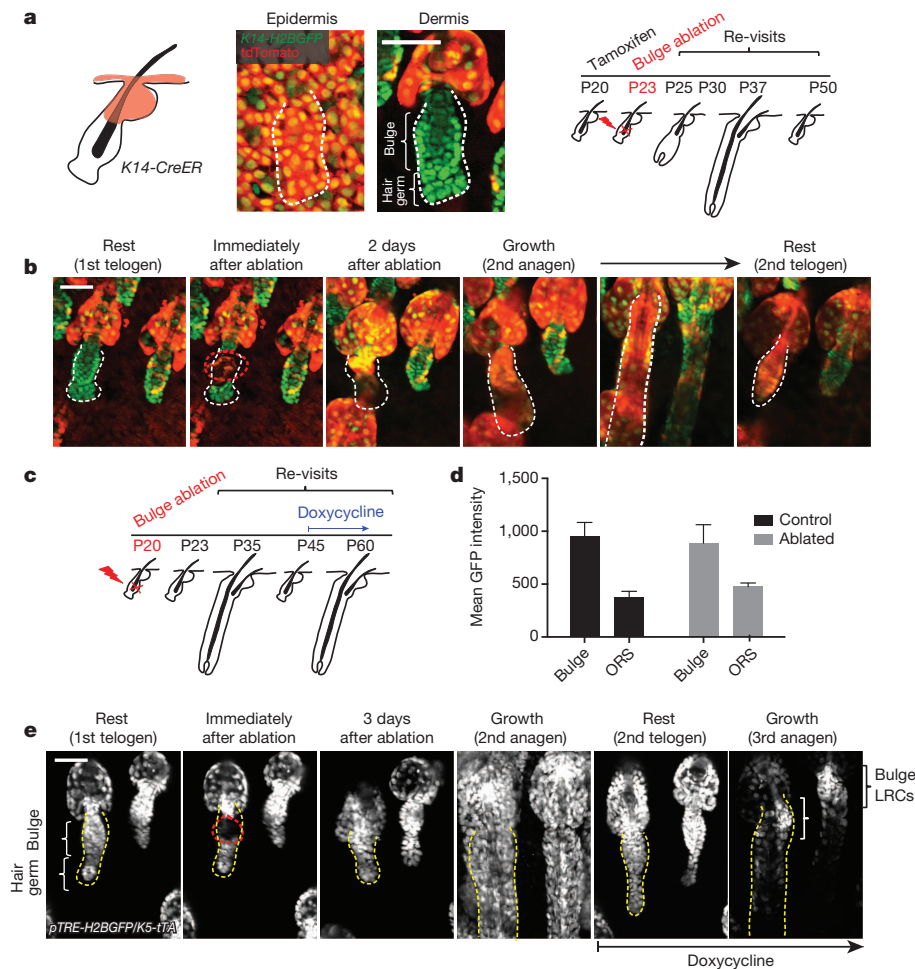


Figure 4 | Functional reconstitution of the stem-cell niche from non-hair epithelial populations. **a**, Scheme of *in vivo* lineage tracing of non-hair epithelial populations after bulge laser ablation. **b**, Example of bulge-ablated hair follicle showing the influx of labelled non-hair epithelial cell populations into the niche and their contribution to hair growth. **c**, Scheme of label

retention experiment after bulge ablation. **d**, Quantification of label retention in control and bulge-ablated hair follicles ($n = 15$ and 12 follicles; error bars represent s.e.m.). **e**, Example of label retention in the niche after bulge ablation and recovery. Scale bars, $50\ \mu\text{m}$.

that part of the niche these cells receive different stimuli pushing them to commit towards a differentiation pathway to support the subsequent hair cycle.

Our model is consistent with previous data from the hair follicle and other stem-cell niches^{22,31,35,36,44,45} but directly demonstrates the significance of the niche for stem-cell fate determination. Our results from the laser ablation experiments further support this notion, highlighting the fact that niche stem cells can be dispensable for tissue regeneration, provided that the overall integrity of the niche is maintained. In this context, injury can induce cell mobility between different tissue compartments but the overall structure and function of the tissue is maintained because cells are capable of adopting new fates as dictated by their new niche microenvironment (Extended Data Fig. 9). This may also explain how certain hierarchies that exist between different stem-cell pools under homeostatic conditions can be re-shuffled and new ones established after injury, as part of a wound healing process⁴⁶. Identifying the extrinsic factors that make up a particular niche microenvironment is paramount for understanding the mechanism of stem-cell fate determination and our ability to manipulate stem cells for therapeutic purposes.

METHODS SUMMARY

K19-CreER mice were created and obtained from G. Gu's laboratory³⁸. *Lgr5-CreER* mice were created by H. Clevers's laboratory¹⁶ and obtained from The Jackson Laboratory. *K14-CreER* mice were created by E. Fuchs's laboratory⁴⁷ and obtained

from The Jackson Laboratory. *Rosa-stop-tdTomato* mice were created by H. Zeng's laboratory³⁹ and obtained from The Jackson Laboratory. Mice were bred to a mixed albino background and males were preferentially used for experiments. All studies and procedures involving animal subjects were approved by the Institutional Animal Care and Use Committee at Yale School of Medicine and conducted in accordance with the approved animal handling protocol. Expression of the Cre fluorescent reporter for the lineage tracing experiments was induced with a single intraperitoneal injection of tamoxifen ($20\ \mu\text{g g}^{-1}$ and $1\ \mu\text{g g}^{-1}$ in corn oil for *K19-CreER* and *Lgr5-CreER*, respectively) at $\sim\text{P20}$ or times specified. For lineage tracing of epithelial populations above the hair follicle, *K14-CreER/Rosa-stop-tdTomato* mice were given a single intraperitoneal injection of tamoxifen ($0.2\ \text{mg g}^{-1}$ in corn oil). For the label retention experiment, *K5-tTA/pTRE-H2BGFP* mice were given doxycycline ($1\ \text{mg ml}^{-1}$) in potable water at times specified. Intravital microscopy and laser ablation procedures were carried out as described previously²⁴. All lineage tracing and ablation experiments were repeated at least in triplicates or otherwise indicated.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 July; accepted 23 August 2013.

Published online 6 October 2013.

- Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441**, 1075–1079 (2006).
- Spradling, A. C. *et al.* Stem cells and their niches: integrated units that maintain *Drosophila* tissues. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 49–57 (2008).

3. Fuchs, E. The tortoise and the hair: slow-cycling cells in the stem cell race. *Cell* **137**, 811–819 (2009).
4. Li, L. & Clevers, H. Coexistence of quiescent and active adult stem cells in mammals. *Science* **327**, 542–545 (2010).
5. Greco, V. & Guo, S. Compartmentalized organization: a common and required feature of stem cell niches? *Development* **137**, 1586–1594 (2010).
6. Copley, M. R., Beer, P. A. & Eaves, C. J. Hematopoietic stem cell heterogeneity takes center stage. *Cell Stem Cell* **10**, 690–697 (2012).
7. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
8. Sato, T. *et al.* Interferon regulatory factor-2 protects quiescent hematopoietic stem cells from type I interferon-dependent exhaustion. *Nature Med.* **15**, 696–700 (2009).
9. Kopp, H.-G., Avecilla, S. T., Hooper, A. T. & Rafii, S. The bone marrow vascular niche: home of HSC differentiation and mobilization. *Physiology (Bethesda)* **20**, 349–356 (2005).
10. Celso, C. L. *et al.* Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. *Nature* **457**, 92–96 (2009).
11. Xie, Y. *et al.* Detection of functional haematopoietic stem cell niche using real-time imaging. *Nature* **457**, 97–101 (2009).
12. Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495**, 231–235 (2013).
13. Greenbaum, A. *et al.* CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. *Nature* **495**, 227–230 (2013).
14. Sangiorgi, E. & Capecchi, M. R. *Bmi1* is expressed *in vivo* in intestinal stem cells. *Nature Genet.* **40**, 915–920 (2008).
15. Takeda, N. *et al.* Interconversion between intestinal stem cell populations in distinct niches. *Science* **334**, 1420–1424 (2011).
16. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
17. Buczak, S. J. A. *et al.* Intestinal label-retaining cells are secretory precursors expressing *Lgr5*. *Nature* **495**, 65–69 (2013).
18. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).
19. Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing *Lgr5* stem cells. *Cell* **143**, 134–144 (2010).
20. Cotsarelis, G., Sun, T. T. & Lavker, R. M. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329–1337 (1990).
21. Tumber, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004).
22. Greco, V. *et al.* A two-step mechanism for stem cell activation during hair regeneration. *Cell Stem Cell* **4**, 155–169 (2009).
23. Ito, M., Kizawa, K., Hamada, K. & Cotsarelis, G. Hair follicle stem cells in the lower bulge form the secondary germ, a biochemically distinct but functionally equivalent progenitor cell population, at the termination of catagen. *Differentiation* **72**, 548–557 (2004).
24. Rompolas, P. *et al.* Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration. *Nature* **487**, 496–499 (2012).
25. Trempus, C. S. *et al.* Enrichment for living murine keratinocytes from the hair follicle bulge with the cell surface marker CD34. *J. Invest. Dermatol.* **120**, 501–511 (2003).
26. Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. & Fuchs, E. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635–648 (2004).
27. Claudinot, S., Nicolas, M., Oshima, H., Rochat, A. & Barrandon, Y. Long-term renewal of hair follicles from clonogenic multipotent stem cells. *Proc. Natl Acad. Sci. USA* **102**, 14677–14682 (2005).
28. Liu, Y., Lyle, S., Yang, Z. & Cotsarelis, G. Keratin 15 promoter targets putative epithelial stem cells in the hair follicle bulge. *J. Invest. Dermatol.* **121**, 963–968 (2003).
29. Ito, M. *et al.* Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nature Med.* **11**, 1351–1354 (2005).
30. Morris, R. J. *et al.* Capturing and profiling adult hair follicle stem cells. *Nature Biotechnol.* **22**, 411–417 (2004).
31. Zhang, Y. V., Cheong, J., Ciapurin, N., McDermitt, D. J. & Tumber, T. Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells. *Cell Stem Cell* **5**, 267–278 (2009).
32. Müller-Röver, S. *et al.* A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages. *J. Invest. Dermatol.* **117**, 3–15 (2001).
33. Jahoda, C. A., Horne, K. A. & Oliver, R. F. Induction of hair growth by implantation of cultured dermal papilla cells. *Nature* **311**, 560–562 (1984).
34. Kaufman, C. K. *et al.* GATA-3: an unexpected regulator of cell lineage determination in skin. *Genes Dev.* **17**, 2108–2122 (2003).
35. Legué, E. & Nicolas, J.-F. Hair follicle renewal: organization of stem cells in the matrix and the role of stereotyped lineages and behaviors. *Development* **132**, 4143–4154 (2005).
36. Hsu, Y.-C., Pasolli, H. A. & Fuchs, E. Dynamics between stem cells, niche, and progeny in the hair follicle. *Cell* **144**, 92–105 (2011).
37. Sequeira, I. & Nicolas, J.-F. Redefining the structure of the hair follicle by 3D clonal analysis. *Development* **139**, 3741–3751 (2012).
38. Means, A. L., Xu, Y., Zhao, A., Ray, K. C. & Gu, G. A. CK19^{CreERT} knockin mouse line allows for conditional DNA recombination in epithelial cells in multiple endodermal organs. *Genesis* **46**, 318–323 (2008).
39. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
40. Youssef, K. K. *et al.* Identification of the cell lineage at the origin of basal cell carcinoma. *Nature Cell Biol.* **12**, 299–305 (2010).
41. Jaks, V. *et al.* *Lgr5* marks cycling, yet long-lived, hair follicle stem cells. *Nature Genet.* **40**, 1291–1299 (2008).
42. Chi, W., Wu, E. & Morgan, B. A. Dermal papilla cell number specifies hair size, shape and cycling and its reduction causes follicular decline. *Development* **140**, 1676–1683 (2013).
43. Van Keymeulen, A. & Blanpain, C. Tracing epithelial stem cells during development, homeostasis, and repair. *J. Cell Biol.* **197**, 575–584 (2012).
44. Barroca, V. *et al.* Mouse differentiating spermatogonia can generate germinal stem cells *in vivo*. *Nature Cell Biol.* **11**, 190–196 (2009).
45. Nystul, T. & Spradling, A. An epithelial niche in the *Drosophila* ovary undergoes long-range stem cell replacement. *Cell Stem Cell* **1**, 277–285 (2007).
46. Plikus, M. V. *et al.* Epithelial stem cells and implications for wound repair. *Semin. Cell Dev. Biol.* **23**, 946–953 (2012).
47. Vasioukhin, V., Degenstein, L., Wise, B. & Fuchs, E. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl Acad. Sci. USA* **96**, 8551–8556 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to S. King, S. Guo and A. Horwich for critical feedback on the manuscript. We thank E. Fuchs for the *K14-H2BGFP*, *Lef1-RFP* and *pTRE-H2BGFP* mice, A. Glick for the *K5-tTA* mice and G. Gu for the *K19-CreER* mice. We thank D. Gonzalez and A. Haberman for technical support with intravital microscopy. P.R. is a New York Stem Cell Foundation–Druckenmiller Fellow. This work was supported by The New York Stem Cell Foundation and by grants to V.G. from the American Cancer Society (RGS-12-059-01-DCC) and the National Institute of Arthritis and Musculoskeletal and Skin Diseases (1R01AR063663-01).

Author Contributions P.R. and V.G. designed experiments and wrote the manuscript; P.R. performed the experiments and analysed the data; K.R.M. assisted with the revisions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.G. (valentina.greco@yale.edu).

The structure of the box C/D enzyme reveals regulation of RNA methylation

Audrone Lapinaite¹, Bernd Simon¹, Lars Skjaerven¹, Magdalena Rakwalska-Bange¹, Frank Gabel^{2,3,4,5} & Teresa Carlomagno¹

Post-transcriptional modifications are essential to the cell life cycle, as they affect both pre-ribosomal RNA processing and ribosome assembly. The box C/D ribonucleoprotein enzyme that methylates ribosomal RNA at the 2'-O-ribose uses a multitude of guide RNAs as templates for the recognition of rRNA target sites. Two methylation guide sequences are combined on each guide RNA, the significance of which has remained unclear. Here we use a powerful combination of NMR spectroscopy and small-angle neutron scattering to solve the structure of the 390 kDa archaeal RNP enzyme bound to substrate RNA. We show that the two methylation guide sequences are located in different environments in the complex and that the methylation of physiological substrates targeted by the same guide RNA occurs sequentially. This structure provides a means for differential control of methylation levels at the two sites and at the same time offers an unexpected regulatory mechanism for rRNA folding.

During the biosynthesis and processing of the pre-rRNA transcripts post-transcriptional modifications of ribonucleotides occur in functional regions, including intersubunit interfaces, decoding and peptidyltransferase centres¹. Among the possible modifications, 2'-O-ribose methylation was shown to protect RNA from ribonucleolytic cleavage², stabilize single base pairs, serve as chaperone^{3,4} and affect folding at high temperatures⁵. rRNA methylation is essential for both pre-rRNA processing and ribosome assembly, with complete suppression of methylation leading to cell death^{6,7}.

In eukaryotes, 2'-O-ribose methylation is carried out by the box C/D small nucleolar RNA-protein complex (snoRNP). The archaeal equivalent box C/D sRNP consists of three core proteins (Extended Data Fig. 1) assembled around the guide small RNA. This RNA contains so-called box C/D and C'/D' motifs, which fold in the K-turn⁸ and K-loop^{8,9} structures, respectively (Extended Data Fig. 1d). Upon substrate binding, the guide sRNA pairs with two different substrate RNAs and selects the methylation site, which is the fifth nucleotide upstream of box D (or D')¹⁰. Using a variety of guide sRNAs, the box C/D snoRNP can methylate more than a hundred different rRNA sequences in humans¹.

In archaea, the assembly of the box C/D sRNP complex is initiated by binding of the L7Ae protein to the K-turn and K-loop elements¹¹. The Nop5-carboxy-terminal domain (CTD) recognizes a composite surface of the L7Ae-sRNA complex, consisting of both RNA and protein^{12,13}. The Nop5-coiled-coil (CC) domain is responsible for self-dimerization, whereas the amino-terminal domain (NTD) interacts with the methylation enzyme, fibrillarin¹⁴, which uses the cofactor S-adenosyl-L-methionine (SAM) as methyl group donor¹⁵.

Recently, the sRNP has been crystallized using an artificial sRNA consisting of two separate strands, base-paired with the corresponding substrates (Extended Data Fig. 1f)¹⁶. In this form, the complex contains two copies of each protein and one sRNA copy (mono-RNP). Despite showing fibrillarin bound to substrate RNA, the structure does not provide a rationale either for the presence of two different guide sequences in all physiological sRNAs or for the asymmetry of the two protein assembly sites (box C/D and C'/D'). Another study of the

sRNP in the absence of substrate RNA used negative stain electron microscopy to derive a different model, where four copies of each protein bind two copies of sRNA¹⁷ (Extended Data Fig. 1g, di-RNP). The position of the RNA could not be determined experimentally; however, in the model proposed to fit the electron microscopy map the two sRNA molecules are located at two opposite sides of a square-shaped complex and fibrillarin is distant from the sRNA, in the so-called 'off' position. In favour of the di-RNP model, recent work has demonstrated that sRNPs assembled around 'one-piece' physiological sRNAs are predominantly di-RNPs, whereas the occurrence of mono-RNPs is significant only when using artificial 'two-pieces' sRNAs¹⁸.

To shed light on the architecture and mechanism of this important enzyme, we study the structure and function of the box C/D sRNP complex from *Pyrococcus furiosus* in solution with a combination of solution-state NMR and small-angle X-ray and neutron scattering (SAXS and SANS). Our study uses a close-to-physiological 'one-piece' sRNA construct and endorses the di-RNP notion. Structures of the apo- and holoenzymes, both of which differ from the ones reported previously^{16,17}, and NMR-based turnover experiments using a physiological guide sRNA allow us to explain the mechanism of rRNA methylation and reveal the sequential regulation of substrate D and D' methylation.

The functional sRNP is dimeric

We reconstituted the box C/D sRNP from full-length recombinant *P. furiosus* L7Ae, Nop5 and fibrillarin and the sRNA of Fig. 1. To simplify the analysis of the NMR spectra, we designed a partially symmetric sRNA (ssR26) starting from the *P. furiosus* sR26 RNA (asR26, Extended Data Fig. 1e), where the K-loop element is substituted by the K-turn element and the two guide sequences are made the same. Importantly, the presence of the apical loop is preserved, thereby preventing full symmetry. The oligomerization state of the box C/D sRNP complex including ssR26 is identical to that of the complex assembled with asR26. In size-exclusion chromatography, the complex elutes as a single peak, corresponding to a di-RNP (~400 kDa) (Extended Data Fig. 1h). To confirm this oligomeric state, we measured

¹European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. ²Université Grenoble Alpes, Institut de Biologie Structurale, F-38027 Grenoble, France. ³Commissariat à l'Energie Atomique, Direction de Sciences du Vivant, Institut de Biologie Structurale, F-38027 Grenoble, France. ⁴Centre National de la Recherche Scientifique, Institut de Biologie Structurale, F-38027 Grenoble, France. ⁵Institut Laue-Langevin, F-38042 Grenoble, France.

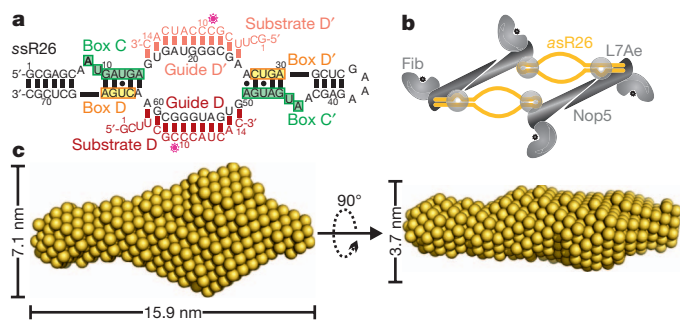


Figure 1 | ssR26 RNA in the *P. furiosus* box C/D complex. **a**, ssR26 RNA used in the sRNP for NMR and SAS experiments. Substrates D (red) and D' (salmon) have the same sequence. A star marks the 2'-O-methylation site. **b**, Scheme of the sRNP explaining the concept of contrast matching in SANS. The data are collected in experimental conditions where the scattering intensity of the [^1H]proteins is masked by matching with the solvent (42%/58% $\text{D}_2\text{O}/\text{H}_2\text{O}$), whereas the [^2H]RNA scattering dominates the curve. Fib, fibrillarin. **c**, *Ab initio* modelling of the ssR26 RNA in the context of the sRNP from SANS data collected as described in **b**. The length of 15.9 nm is considerably larger than the 11 nm expected for the mono-RNP model¹⁶ and accommodates two ssR26 molecules.

the SAXS scattering profile of the particle. The radius of gyration of $58.2 \pm 0.6 \text{ \AA}$ (Extended Data Table 1) is significantly larger than that predicted from the crystallized mono-RNP (40.0 \AA)¹⁶. Last, we performed native gel electrophoresis, which also confirmed that the complex has a molecular weight of $\sim 400 \text{ kDa}$ (Extended Data Fig. 1i).

Two fibrillarin settings in the apo-sRNP

Next, we determined the structure of the 390 kDa box C/D complex in the absence of substrate RNA by a combination of NMR and SAXS/SANS data (Fig. 2). We reasoned that the structures of L7Ae and fibrillarin, as well as those of the Nop5 domains and the sRNA K-turn modules, do not change with respect to their structures in the L7Ae-K-turn-RNA¹⁹ and Nop5-fibrillarin¹⁴ complexes determined previously. In addition, both the interaction surfaces of the Nop5-CTD with the L7Ae-K-turn-RNA complex^{12,13} and of the Nop5-NTD with fibrillarin in the Nop5-fibrillarin¹² complex are likely to be conserved in the full enzyme. To verify these assumptions, we specifically labelled the methyl groups of Ile, Val and Leu of fibrillarin and L7Ae²⁰ with ^{13}C and ^1H and monitored the chemical-shift perturbations upon stepwise formation of the complex by methyl-transverse relaxation-optimized spectroscopy experiments. The NMR signals of methyl groups can also be observed for complexes as large as several hundreds of kDa, owing to

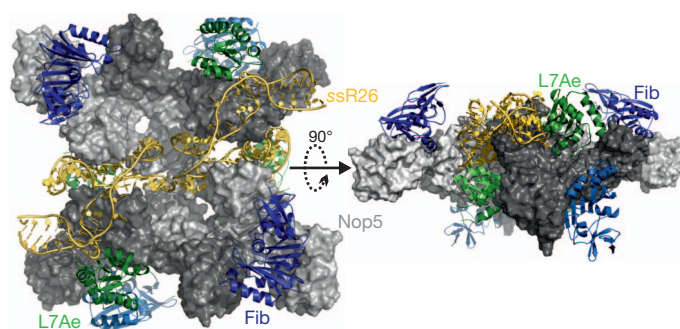


Figure 2 | Non-equivalent fibrillarin environments in the apo-box C/D sRNP. Structure of the apo-box C/D sRNP. The Nop5 copies (two associated with the box C/D (C'/D') elements, dark (light) grey, numbering in Extended Data Fig. 6) form a platform. Two fibrillarin copies (associated with the box C'/D' elements, dark blue) are above the Nop5 platform, on the same side as the sRNA; two other copies (associated with the box C/D elements, light blue) are below the Nop5 platform, on the opposite side as the sRNA. The two L7Ae associated with the box C/D (C'/D') elements are in dark (light) green.

their favourable relaxation properties and the high intensity of the signal generated by three protons. In addition to chemical-shift perturbations data, we also collected paramagnetic relaxation enhancement (PRE) data. In this experiment, a tag carrying a free electron is coupled with a single cysteine engineered, for example, in Nop5. The relaxation enhancements elicited on the methyl groups of either L7Ae or fibrillarin by the unpaired electron in Nop5 are quantified and translated into distances²¹, which can be used to determine the mutual orientation of the interacting protein domains. All experiments confirmed that the previously described interfaces are preserved in the full complex (Extended Data Fig. 2).

On the basis of this analysis, we built the structure of the box C/D sRNP using the conformations of the modules L7Ae-K-turn-RNA-Nop5-CTD and Nop5-NTD-fibrillarin observed in previous crystal structures, and restricted our conformational search to the orientation of the three domains of Nop5, the conformation of the sRNA in parts other than the K-turn motifs and A-form helices and the relative position of the four copies of proteins in the di-RNP complex (Extended Data Fig. 3). To this end, we designed a structure calculation protocol that capitalizes on the following experimental data: (1) methyl resonances chemical-shift perturbations of L7Ae and fibrillarin upon formation of the full complex; (2) PRE data defining the relative orientation of the pairs fibrillarin-L7Ae, fibrillarin-Nop5-CTD, L7Ae-Nop5-NTD, fibrillarin-Nop5-CC and L7Ae-Nop5-CC; (3) contrast matching SANS data defining the individual shapes of each component in the context of the full complex (Extended Data Fig. 4).

The $P(r)$ distribution of the complex containing either ^2H -labelled fibrillarin (Fig. 3b) or L7Ae confirms that the particle contains more than two copies of each protein. Unexpectedly, the SANS scattering curve of the [^2H]RNA complex, acquired in a 58%/42% $\text{H}_2\text{O}/\text{D}_2\text{O}$ solution to mask the scattered intensity of the [^1H]proteins, is best fitted by a continuous distribution of atoms with dimensions $15.9 \times 7.1 \times 3.7 \text{ nm}$, which clearly indicates that the two ssR26 copies are close to each other in the complex (Fig. 1c) and disproves the arrangement of the RNA in the electron-microscopy-derived model (Extended Data Figs 1g and 4e)^{12,17}.

The structure calculation protocol used 452 PRE data-points and SAXS/SANS data to derive the structure of the apo-sRNP (without substrate RNAs) in solution at 4.8 \AA precision (Fig. 2, Extended Data Fig. 5a and Extended Data Table 2). The four Nop5-CC domains form a platform, and the two sRNA molecules pack together to yield an elongated shape, which lies on this platform at an angle of about 45° (Fig. 2 and Extended Data Fig. 6). The box C/D elements are found either at the two extremities of the rod-like structure or in the centre. Both RNA conformations are *de facto* equivalent when using ssR26; this identity is lifted in the asR26 RNA, with consequences in the context of the holo-complex, as discussed later. Electrostatic interactions between the RNA guide sequences and the Nop5 protein stabilize the complex structure (Extended Data Fig. 6b) and are in excellent agreement with published cross-link data²². Despite confirming the di-RNP architecture, the *P. furiosus* apo-complex structure does not match the electron microscopy envelopes of box C/D sRNPs from other organisms^{17,18} (Extended Data Fig. 4h).

The four fibrillarin copies do not contact the sRNA, in agreement with the very weak chemical-shift perturbations upon transition from the Nop5-fibrillarin complex to the full sRNP (Extended Data Fig. 7a) and with previous electron-microscopy-guided modelling¹⁷. Instead, they rest at the end of the Nop5-CC domains: two copies are above the platform defined by the Nop5 proteins on the same side as the sRNA, whereas the other two copies are on the opposite side (Fig. 2b). This arrangement predicts that only two fibrillarin copies can reach the guide RNA sequences to yield a methylation-competent conformation.

Both the PRE and the SANS data place the four fibrillarin copies in the 'off' position. However, a small subset of PRE effects (V11, L15, I53, I75, I78), measured for L7Ae in combination with the Nop5(E68C) mutant, cannot be fitted together with the other data; rather, they

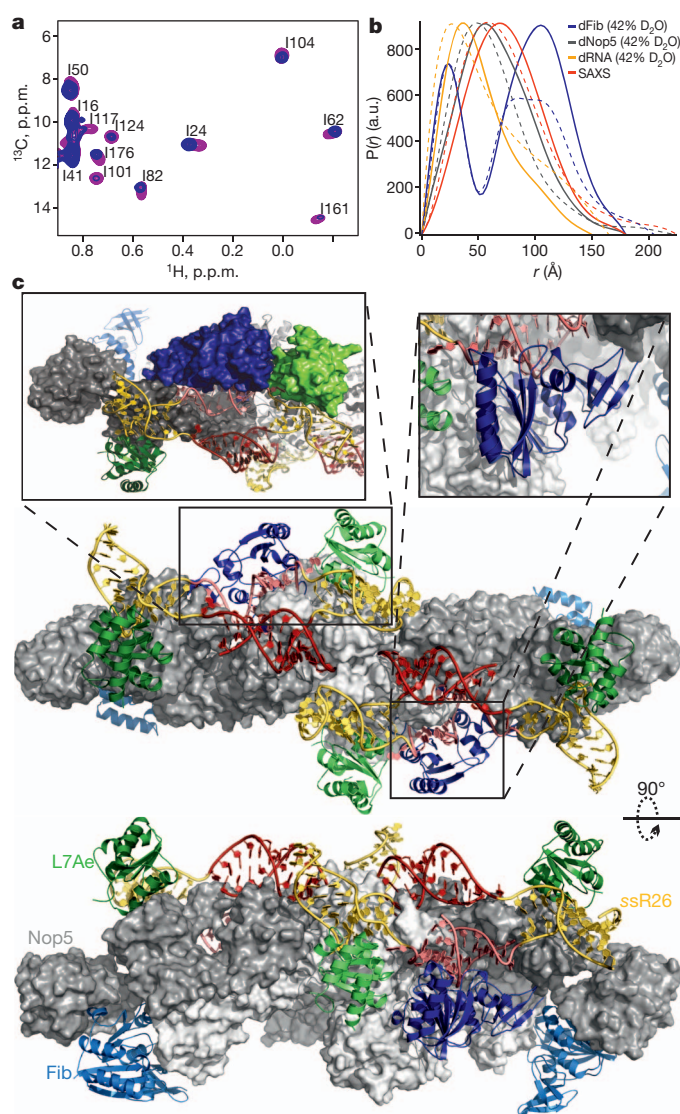


Figure 3 | Catalytic structure of the box C/D sRNP. **a**, Isoleucine region of the ^{13}C - ^1H correlations of the fibrillar ILV-methyl groups in the apo-enzyme (blue) and after addition of substrate RNA (pink). Methyl groups of I24, L58, I62, I82, I117, I176, V196 and L210 split in two sets. **b**, Pair-wise distance distribution $P(r)$, calculated from the scattering curves of the box C/D complex in the absence (solid lines) and in the presence (dashed lines) of substrate RNA and assembled from all non-labelled components (SAXS), ^2H RNA and ^1H proteins (dRNA, SANS), ^2H Nop5, ^1H others (dNop5, SANS), ^2H fibrillar, ^1H others (dFib, SANS). a.u., arbitrary units. **c**, Structure of the holo-box C/D sRNP (colour code as in Fig. 2). Two fibrillar copies (light blue) are in the 'off' position, on the opposite side from the corresponding guide-substrate D duplexes (firebrick). Two fibrillar copies (dark blue) contact the guide-substrate D' duplexes (salmon, right insert) and are able to perform methylation. The fibrillar is directed to the methylation site by packing with L7Ae and Nop5-CTD (left insert).

mimic those obtained later for the holo-complex (Extended Data Fig. 7d), where fibrillar is in contact with the guide-substrate duplex. This observation offers evidence for occasional fly-casting motions of the catalytic module fibrillar that allow the protein to visit the space close to the sRNA in the apo-complex and probably aid the recognition of the guide-substrate duplex.

Structural change upon substrate binding

To investigate the structure of the complex with bound substrate (holo-form), we titrated a 16-nt single-stranded RNA, complementary to the guide sequence, and monitored the particle by both NMR

and SAXS/SANS. In the ^{13}C - ^1H correlation of Ile-Leu-Val (ILV)-methyl labelled fibrillar, a few resonances divide into two peaks (Fig. 3a), one of which conserves the same position as in the spectrum of the apo-complex. The intensity ratio of the two peaks increases upon addition of substrate RNA and plateaus at 1:1 in the presence of excess substrate. This observation demonstrates that in the holo-complex only two of the four fibrillar copies are in contact with the substrate RNAs, whereas the other two copies occupy the same environment as in the apo-complex. In contrast, all four guide sequences base pair with the substrate or product RNA, as indicated by the absence of free RNA resonances upon addition of a stoichiometric amount of either non-methylated or methylated substrate (data not shown).

The SAXS/SANS $P(r)$ functions of the ^1H all, ^2H Nop5 and ^2H fibrillar complexes reveal a large conformational rearrangement upon substrate binding (Fig. 3b). These curves were used, together with 257 PRE distances, to obtain the structure of the holo-sRNP at 5.2 Å precision (Fig. 3c, Extended Data Figs 5 and 8 and Extended Data Table 2). In agreement with the chemical-shift perturbations data, we also imposed interaction restraints between the SAM binding site of two fibrillar copies and the two accessible guide-substrate RNA duplexes.

Upon binding of the substrate RNA, the ssR26 molecules transition from a bent to an elongated form (Fig. 3c and Supplementary Video 1), to allow for the formation of the A-form helices between the guide and the substrate sequences. Pulled by the RNA, the two Nop5 dimers move apart from each other, generating an elongated shape (Fig. 3b). Both D and D' guide-substrate helices are roughly parallel to the stems flanking box C/D and C'/D'. The two D' guide-substrate helices are on the same plane as the Nop5 proteins and can be contacted by fibrillar (dark blue, Fig. 3c); the D guide-substrate helices are above the plane of the Nop5 proteins, on the opposite side as the other two fibrillar molecules (light blue, Fig. 3c).

Upon elongation of the complex, two Nop5-CTDs move to the centre of the sRNP and contact each other through helices $\alpha 6$, $\alpha 10$ and $\alpha 12$, the tip of $\alpha 7$ and loop $\alpha 10$ - $\alpha 11$ (Extended Data Fig. 8b). These Nop5-CTD elements form a composite surface, which can associate with itself through complementarities in electrostatics and shape.

The fibrillar molecule that reaches substrate D' is the one connected to the Nop5 protein bound to the box C'/D' of the other sRNA

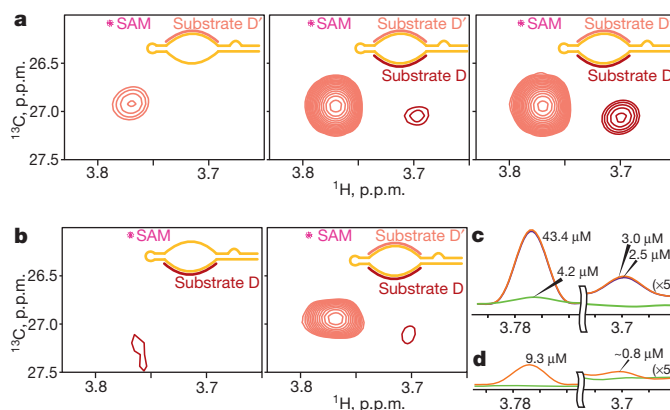


Figure 4 | Methylation of substrate D and D' is sequentially regulated.

The release of ^{13}C -labelled 2'-OCH₃ substrate RNA is monitored in ^{13}C - ^1H correlations (22 h) after addition of ^{13}C -labelled SAM and excess of substrate RNA. The peak at 57.9 (folded at 26.8, ^{13}C) and 3.77 (^1H) p.p.m. corresponds to the methyl group of product D', that at 58.1 (folded at 27.0, ^{13}C) and 3.69 (^1H) p.p.m. to the methyl group of product D. Total [sRNA], 15 μM . **a**, Left, addition of substrate D' (57 μM) and SAM (114 μM) (green trace in c); centre, addition of substrate D (57 μM) and SAM (114 μM) (purple trace in c); right, addition of SAM (114 μM) (orange in c). **b**, Left, addition of substrate D (42 μM) and SAM (84 μM) (green trace in d); centre, addition of substrate D' (42 μM) and SAM (84 μM) (orange in d). **c**, **d**, Time course of **a**, **b** with quantification of the amount of released product. Traces were taken across the two-dimensional peak of the corresponding species.

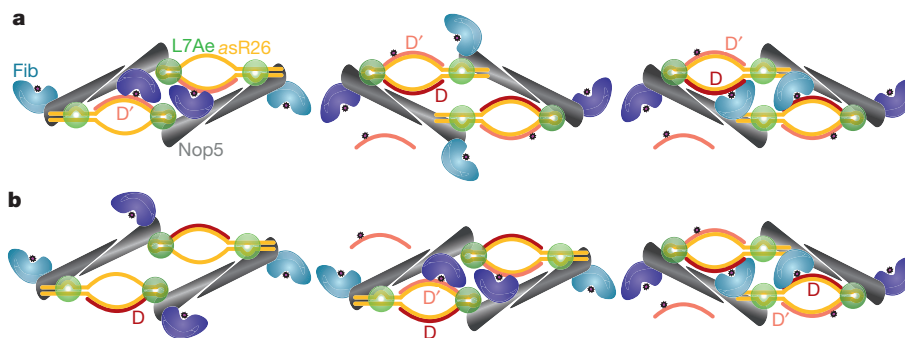


Figure 5 | Regulatory mechanism for sequential methylation.

Schematically, the substrate sequences depicted in the inside of the complex are reached by the methyltransferase for methylation, whereas the external ones are far from all fibrillar copies. In the apo-complex the guide D' (D) sequences are close to (far from) fibrillar. **a**, Left, substrate D' binds first on the same plane as two fibrillar copies and is methylated (Fig. 4a, left); subsequent addition of substrate D causes a conformational switch that brings the

guide-substrate D duplexes and two fibrillar copies on one plane (centre). Right, substrate D is methylated (Fig. 4a, right). **b**, Left, substrate D, added first, binds to the guide D sequences far from fibrillar, where it cannot be methylated (Fig. 4b, left); substrate D', added next, binds to the guide D' sequences next to two fibrillar copies and is methylated. Right, the conformational switch takes place and substrate D is methylated as well (Fig. 4b, centre).

molecule (Extended Data Fig. 8). In this respect, the recognition of the guide-substrate duplex occurs *in trans*. Packing with the CTD of the Nop5 copy bound to the box C/D of the same sRNA molecule and the L7Ae at the box C'/D' also of the same sRNA molecule (Fig. 3c and Extended Data Fig. 8c, d) confines fibrillar to a restricted space on the substrate RNA and promotes specific methylation at position -5 from box D'. In particular, the N-terminal tip of fibrillar interacts with the Nop5-CTD through charge complementarities (Extended Data Fig. 8d). This tip is not part of the methyltransferase domain (beyond amino acid position 50) and its function has remained obscure; our structure reveals that it is indispensable for site-specific methylation.

Binding of fibrillar to the guide-substrate duplex is dependent on the presence of a free 2'-OH: when C9-2'-OCH₃ substrate RNA is added to the apo-sRNP, the fibrillar ¹³C-¹H correlation does not change (Extended Data Fig. 7c), indicating that the protein does not bind the methylated duplex. The difference in fibrillar affinity for methylated versus non-methylated duplex is likely to trigger the release of the methyltransferase from the product, thereby initiating turnover.

Controlled, site-specific methylation

To prove that the di-sRNP carries out sequence-specific methylation, we developed an NMR-based assay to observe enzyme activity in the same conditions as those used for the structural studies: the release of [¹³C]methyl-O-RNA is monitored in a ¹³C-¹H correlation experiment after addition of ¹³C-labelled SAM and excess of substrate RNA. The presence of a single methyl resonance in the expected region confirms that the methylation occurs at a specific position, which, by comparison with the NMR spectrum of synthetic ¹³C-labelled product D', can be identified as the fifth nucleotide upstream of box D'. This assay confirms in a unique way that the determined structure corresponds to the active form of the enzyme.

The most interesting feature of our structure is that the binding environments of substrate D and D' are not equivalent. Consequently, the guide-substrate duplexes (D) located above the proteins plane (Fig. 3c) are not accessible to fibrillar and would have to exchange places with the other two duplexes (D') to become modified. To corroborate this hypothesis, we assembled the complex with the physiological asR26 RNA (Extended Data Fig. 1e) and monitored by NMR the binding, methylation and release of substrate D and D' upon addition of ¹³C-labelled SAM. Addition of substrate D' to the apo-complex yields the release of product D' (Fig. 4a). Subsequent addition of substrate D increases the turnover for substrate D' and at the same time results in the release of smaller amounts of product D. In the context of our holoenzyme structure, the higher turnover rate in the presence of both substrates can be attributed to the additional strain on

the RNA imposed by formation of all four guide-substrate helices, which likely induces the separation of the products after fibrillar has separated from them. Intriguingly, when substrate D is added first, we do not detect any product D. Further addition of substrate D' yields methylation and release of product D', whereas methylated substrate D is detected only at this point and in small amounts (Fig. 4b). This observation, together with the structure of the holoenzyme, suggests that methylation at the two sites occurs sequentially in a well-defined order; for the sR26 RNA, methylation of substrate D' precedes modification of substrate D. Furthermore, a higher turnover is measured for substrate D' than for substrate D under our experimental conditions. Such sequential control of site-specific methylation serves the purpose of avoiding methylation of the D site in the absence of methylation at the D' site, explains why different guide sequences are combined on the same guide sRNA and suggests that their combination is not casual.

Discussion

To visualize the mechanism of sequential methylation in structural terms we propose the following model (Fig. 5). In the absence of any substrate, the asymmetric sRNP preferentially assumes a conformation where the fibrillar associated with the box C'/D' is on the same plane as the RNA (Fig. 5). In this conformation, substrate D' can be loaded on the complex, methylated and released. Addition of the other substrate to the complex half-loaded with product RNA triggers a conformational switch, whereby the newly formed duplexes and the corresponding fibrillar copies move to one plane, allowing the methylation of substrate D. On the other hand, binding of substrate D in the absence of substrate D' to the sites that are far from fibrillar does not lead to product D (Fig. 5b). *In vivo*, the sRNP is probably recycled by removal of all products after modification of all four sites.

Based on our data we propose that rRNA methylation by the sR26 guide RNA occurs in a sequential manner, providing a novel mechanism to regulate correct establishment of rRNA methylation patterns. Although the role of sequential methylation remains to be confirmed for other guide sRNAs, it is tempting to speculate that a regulated order of methylation at different rRNA sites may offer elegant means to control the pathway of rRNA folding, a complex process that in part takes place contemporarily to pre-rRNA modification.

METHODS SUMMARY

His₆-tagged proteins were expressed in *Escherichia coli* and purified by affinity chromatography. The RNA was obtained by *in vitro* transcription. The complexes were assembled stepwise and purified by size-exclusion chromatography. SAXS and SANS data were collected for 26 samples with different uniformly deuterated components in buffers containing 0%, 42% or 70% D₂O. Complex concentrations varied between 1.5 and 50 μM.

NMR experiments were performed on Bruker Avance III 600 MHz and 800 MHz spectrometers, equipped with HCN triple-resonance cryo-probes at 328 K. L7Ae and fibrillarin were deuterated and specifically ^{13}C -, ^1H -labelled at the ILV-methyl groups for NMR studies of the full sRNP. Sample concentrations were between 7 and 40 μM . Paramagnetic relaxation enhancement (PRE) was applied to measure inter-protein distances (up to 25 Å) within the apo- and holo-sRNPs. Structures were calculated by a simulated annealing protocol with fixed structural domains connected by flexible hinge regions. The protein starting structures were generated by reading the X-ray coordinates of each protein component from 3NMU.pdb. The PRE data were entered as distances restraints. In addition, a dummy atom representation of the SANS-derived low-resolution shape of the RNA was used with ambiguous distance restraints between each dummy atoms and all P, C1' and C4 of the RNA and vice-versa. The final structures were selected by the requirement to fit all SAXS/SANS curves and the experimental PRE distances simultaneously.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 June; accepted 19 August 2013.

Published online 13 October 2013.

- Decatur, W. A. & Fournier, M. J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**, 344–351 (2002).
- Herschlag, D., Eckstein, F. & Cech, T. R. The importance of being ribose at the cleavage site in the *Tetrahymena* ribozyme reaction. *Biochemistry* **32**, 8312–8321 (1993).
- Williams, D. J., Boots, J. L. & Hall, K. B. Thermodynamics of 2'-ribose substitutions in UUCG tetraloops. *RNA* **7**, 44–53 (2001).
- Helm, M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* **34**, 721–733 (2006).
- Kawai, G. *et al.* Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* **31**, 1040–1046 (1992).
- Schimmang, T., Tollervey, D., Kern, H., Frank, R. & Hurt, E. A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. *EMBO J.* **8**, 4015–4024 (1989).
- Tollervey, D., Lehtonen, H., Carmo-Fonseca, M. & Hurt, E. The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J.* **10**, 573–583 (1991).
- Kiss-László, Z., Henry, Y. & Kiss, T. Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* **17**, 797–807 (1998).
- Nolivos, S., Carpousis, A. J. & Clouet-d'Orval, B. The K-loop, a general feature of the *Pyrococcus* C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res.* **33**, 6507–6514 (2005).
- Reichow, S. L., Hamma, T., Ferre-D'Amare, A. R. & Varani, G. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.* **35**, 1452–1464 (2007).
- Kuhn, J. F., Tran, E. J. & Maxwell, E. S. Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res.* **30**, 931–941 (2002).
- Xue, S. *et al.* Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol. Cell* **39**, 939–949 (2010).
- Liu, S. *et al.* Binding of the human Prp31 nop domain to a composite RNA-protein platform in U4 snRNP. *Science* **316**, 115–120 (2007).
- Aittaleb, M. *et al.* Structure and function of archaeal box C/D sRNP core proteins. *Nature Struct. Mol. Biol.* **10**, 256–263 (2003).
- Wang, H., Boisvert, D., Kim, K. K., Kim, R. & Kim, S. H. Crystal structure of a fibrillarin homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 Å resolution. *EMBO J.* **19**, 317–323 (2000).
- Lin, J. *et al.* Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* **469**, 559–563 (2011).
- Bleichert, F. *et al.* A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science* **325**, 1384–1387 (2009).
- Bower-Phipps, K. R., Taylor, D., Wang, H. & Baserga, S. The box C/D sRNP dimeric architecture is conserved across domain Archaea. *RNA* **18**, 1527–1540 (2012).
- Moore, T., Zhang, Y., Fenley, M. O. & Li, H. Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure* **12**, 807–818 (2004).
- Tugarinov, V. & Kay, L. E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc.* **125**, 13868–13878 (2003).
- Battiste, J. L. & Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry* **39**, 5355–5365 (2000).
- Ghalei, H., Hsiao, H.-H., Urlaub, H., Wahl, M. C. & Watkins, N. J. A novel Nop5-sRNA interaction that is required for efficient archaeal box C/D sRNP formation. *RNA* **16**, 2341–2348 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by DFG grant CA294/3-1, by EU FP7 ITN project RNPnet (contract number 289007) and by the EMBL. For the SAXS and SANS experiments we thank the ESRF and ILL, BAG MX1219 and the D22 BAG system for beam-time as well as A. Martel and P. Pernot for help with the setup of the instruments. We thank J. Wöhnert for the gift of ^{13}C -labelled SAM and C. Kreutz for the gift of $[\text{C}^{13}]\text{OCH}_3$ substrate RNA. We thank John Kirkpatrick for the assignment of the L7Ae resonances.

Author Contributions A.L., B.S., F.G. performed experiments and analysed data; M.R.-B. prepared the samples; L.S. and B.S. performed calculations; T.C. designed the project, analysed data and wrote the manuscript together with A.L.

Author Information Coordinates of the holoenzyme have been deposited in the Protein Data Bank under accession code 4BY9. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.C. (teresa.carlomagno@embl.de).

A galaxy rapidly forming stars 700 million years after the Big Bang at redshift 7.51

S. L. Finkelstein¹, C. Papovich², M. Dickinson³, M. Song¹, V. Tilvi², A. M. Koekemoer⁴, K. D. Finkelstein¹, B. Mobasher⁵, H. C. Ferguson⁴, M. Giavalisco⁶, N. Reddy⁵, M. L. N. Ashby⁷, A. Dekel⁸, G. G. Fazio⁷, A. Fontana⁹, N. A. Grogin⁴, J.-S. Huang⁷, D. Kocevski¹⁰, M. Rafelski¹¹, B. J. Weiner¹² & S. P. Willner⁷

Of several dozen galaxies observed spectroscopically that are candidates for having a redshift (z) in excess of seven, only five have had their redshifts confirmed via Lyman α emission, at $z = 7.008, 7.045, 7.109, 7.213$ and 7.215 (refs 1–4). The small fraction of confirmed galaxies may indicate that the neutral fraction in the intergalactic medium rises quickly at $z > 6.5$, given that Lyman α is resonantly scattered by neutral gas^{3,5–8}. The small samples and limited depth of previous observations, however, makes these conclusions tentative. Here we report a deep near-infrared spectroscopic survey of 43 photometrically-selected galaxies with $z > 6.5$. We detect a near-infrared emission line from only a single galaxy, confirming that some process is making Lyman α difficult to detect. The detected emission line at a wavelength of 1.0343 micrometres is likely to be Lyman α emission, placing this galaxy at a redshift $z = 7.51$, an epoch 700 million years after the Big Bang. This galaxy's colours are consistent with significant metal content, implying that galaxies become enriched rapidly. We calculate a surprisingly high star-formation rate of about 330 solar masses per year, which is more than a factor of 100 greater than that seen in the Milky Way. Such a galaxy is unexpected in a survey of our size⁹, suggesting that the early Universe may harbour a larger number of intense sites of star formation than expected.

We obtained near-infrared (NIR) spectroscopy of galaxies originally discovered in the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS)^{10,11} with the newly commissioned NIR spectrograph MOSFIRE¹² on the Keck I 10-m telescope. From a parent sample of over 100 galaxy candidates at $z > 7$ in the GOODS-North field selected via their Hubble Space Telescope (HST) colours through the photometric redshift technique^{13–16}, we observed 43 candidate high-redshift galaxies over two MOSFIRE pointings with exposure times of 5.6 and 4.5 h, respectively. Our observations covered Lyman α ($\text{Ly}\alpha$) emission at redshifts of 7.0–8.2. We visually inspected the reduced data at the expected slit positions for our 43 observed sources and found plausible emission lines in eight objects, with only one line detected at $>5\sigma$ significance. The detected emission line is at a wavelength of 1.0343 μm with an integrated signal-to-noise ratio of 7.8 (Fig. 1) and comes from the object designated z8_GND_5296 in our sample (right ascension 12 h 36 min 37.90 s, declination $62^\circ 18' 8.5''$, J2000). On the basis of arguments outlined below (and discussed extensively in the Supplementary Information), we identify this line as the $\text{Ly}\alpha$ transition of hydrogen at a line-peak redshift of $z = 7.5078 \pm 0.0004$; this is consistent with our photometric redshift 95% confidence range of $7.3 < z < 8.1$ for z8_GND_5296.

As expected for a galaxy at $z = 7.51$, z8_GND_5296 is undetected in the HST optical bands, including an extremely deep 0.8 μm image (Fig. 2). The galaxy is bright in the HST NIR bands, becoming brighter

with increasing wavelength, implying that the Lyman break lies near 1 μm and that the galaxy has a moderately red rest-frame ultraviolet colour. The galaxy is well-detected in both Spitzer/IRAC bands (3.6 μm and 4.5 μm wavelength) and is much brighter at IRAC 4.5 μm than at IRAC 3.6 μm . The strong break at observed 1 μm restricts the observed emission line to be either $\text{Ly}\alpha$ at $z = 7.51$ (near the Lyman break) or [O II] 3,726 and 3,729 \AA (a doublet) at $z = 1.78$ (near the rest-frame Balmer/4,000 \AA break). We investigated these two possibilities by comparing our observed photometry to a suite of stellar population models at both redshifts (Fig. 3). A much better fit to the data is obtained when

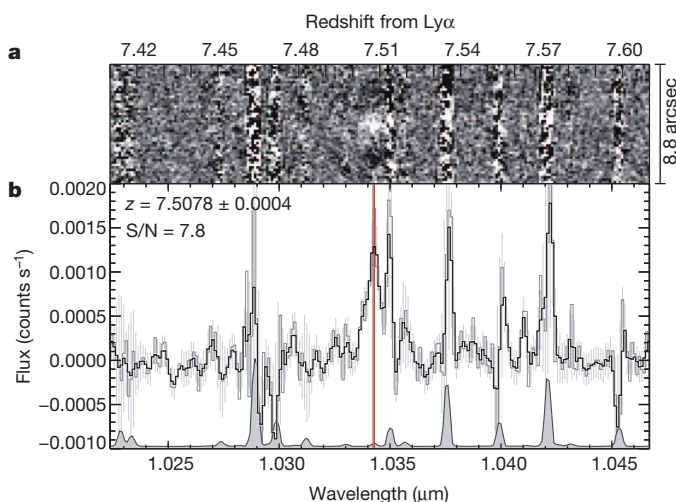


Figure 1 | The observed NIR spectrum of the galaxy z8_GND_5296. **a**, The reduced two-dimensional spectrum. An emission line is clearly seen as a positive signal (white) in the centre, with the negative signals (black) above and below being a result of our ‘dithering’ pattern in the spatial direction along the slit; this is a pattern only exhibited for real objects. **b**, The extracted one-dimensional spectrum (black, smoothed to the spectral resolution; grey, not smoothed). The sky spectrum is shown as the filled grey curve with the scale reduced greatly compared to that of the data. We measure the line to have a signal-to-noise (S/N) of 7.8, and it is also clearly detected in separate reductions of the first and second halves of the data with signal-to-noise ratios of 6.4 and 5.2, respectively. The line has a full-width at half-maximum (FWHM) of 7.7 \AA and is clearly resolved compared to nearby sky emission lines, which have FWHM = 2.7 \AA . The red line denotes the peak flux of the detected emission line, which corresponds to a redshifted $\text{Ly}\alpha$ line at $z = 7.5078 \pm 0.0004$. All other strongly positive or negative features are subtraction residuals due to strong night sky emission. Although the line appears symmetric, there is a sky line residual just to the red of our detected emission line, which makes a measurement of our line’s asymmetry difficult.

¹The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, Texas 78712, USA. ²George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, 4242 TAMU, College Station, Texas 77743, USA. ³National Optical Astronomy Observatory, Tucson, Arizona 85719, USA. ⁴Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. ⁵University of California, Riverside, California 92521, USA. ⁶University of Massachusetts, Amherst, Massachusetts 01003, USA. ⁷Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA. ⁸Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel. ⁹INAF-Osservatorio di Roma, II-00040, Monteporzio, Italy. ¹⁰University of Kentucky, Lexington, Kentucky 40506, USA. ¹¹Infrared Processing and Analysis Center, MS 100-22, Caltech, Pasadena, California 91125, USA. ¹²Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, Arizona 85721, USA.

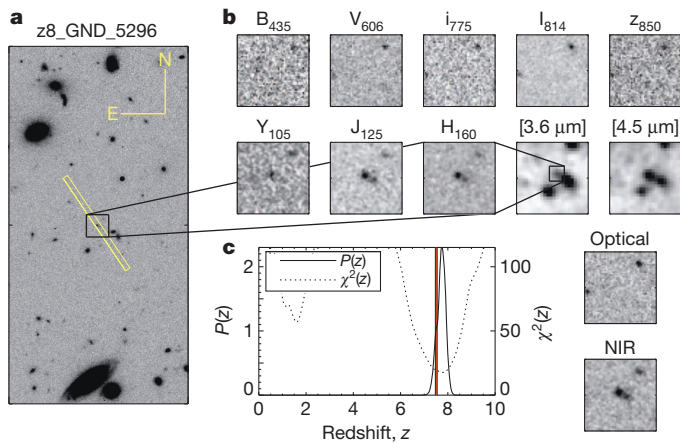


Figure 2 | Images of $z8_GND_5296$. **a**, A portion of the CANDELS/GOODS-N field, shown in the F160W filter (centred at $1.6\ \mu\text{m}$), around $z8_GND_5296$. CANDELS provides the largest survey volume in the distant Universe deep enough to find $z > 7$ galaxies. The $15'' \times 0.7''$ slit is shown as the yellow rectangle. **b**, Magnified multi-wavelength images of the boxed area in **a** (around $z8_GND_5296$); the GOODS and CANDELS HST images (top row, first three images in bottom row) are $3''$ on a side, while the S-CANDELS Spitzer/IRAC 3.6 and $4.5\ \mu\text{m}$ images (last two images in bottom row) are $15''$ on a side. We also show at bottom right mean stacks of the five optical bands and the three NIR bands, the latter showing that this galaxy appears to have a clumpy morphology. This galaxy is not detected in any optical band, even when stacked together, which is strongly suggestive of a redshift greater than 7. The IRAC bands show a faint detection at $3.6\ \mu\text{m}$ and a strong detection at $4.5\ \mu\text{m}$. This signature is expected if strong $[\text{O III}]$ emission is present in the $4.5\text{-}\mu\text{m}$ band, which would be the case for a strongly star-forming galaxy at $z \approx 7.5$ with sub-solar (though still significant) metal content ($0.2\text{--}0.4$ times solar). **c**, The results of our photometric redshift analysis placing $z8_GND_5296$ at $7.3 < z < 8.1$ at 95% confidence, which encompasses our measured spectroscopic redshift (denoted by the vertical line). We show both the probability distribution function as well as the values of χ^2 at each redshift from the photometric redshift analysis; though a low-redshift solution is possible, it is strongly disfavoured, with the high-redshift solution being $\sim 7 \times 10^9$ times more probable.

using models at $z = 7.51$ than at $z = 1.78$, supporting our identification of the emission line as $\text{Ly}\alpha$. Specifically, the model at $z = 1.78$ would result in greater than 4σ significant flux in the $0.8\ \mu\text{m}$ image as well as a near-zero IRAC colour, neither of which is seen. Additionally, the $z = 1.78$ model requires an ageing stellar population with no active star formation, which would have insignificant $[\text{O II}]$ emission. In the Supplementary Information, we discuss a number of tests performed to discern between the $\text{Ly}\alpha$ and $[\text{O II}]$ hypotheses. In summary, although we cannot robustly measure the line asymmetry owing to the nearby sky residual, the spectral energy distribution (SED) fitting results and the lack of a detected second line in the $[\text{O II}]$ doublet lead us to conclude that the detected emission line is $\text{Ly}\alpha$ at $z = 7.51$.

This galaxy is very bright in the rest-frame ultraviolet and optical, with an apparent magnitude of $m_{F160W} = 25.6$ and a derived stellar mass of $1.0^{+0.2}_{-0.1} \times 10^9 M_\odot$ (M_\odot , solar mass). The blue $H - 3.6\ \mu\text{m}$ colour suggests that the moderately red ultraviolet colour ($J - H = 0.1\ \text{mag}$) is due to dust attenuation rather than the intrinsic red colour of an old stellar population. The presence of dust extinction leads to a higher inferred ultraviolet luminosity. To derive the intrinsic star-formation rate (SFR) for this galaxy, we measured a time-averaged SFR from the best-fitting stellar population models to find $\text{SFR} = 330^{+710}_{-10} M_\odot \text{yr}^{-1}$. The very red $3.6\ \mu\text{m} - 4.5\ \mu\text{m}$ colour at $z = 7.51$ can only be due to strong $[\text{O III}]$ 5,007 Å line emission in the $4.5\ \mu\text{m}$ band; indeed, the SED fitting implies an $[\text{O III}]$ 5,007 Å rest-frame equivalent width of 560–640 Å (68% confidence), with a line flux of $5.3 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. This very high $[\text{O III}]$ equivalent width constrains the abundance of metals in this galaxy, as highly enriched stars do not produce hard-enough ionizing spectra, and very low-metallicity systems do not have

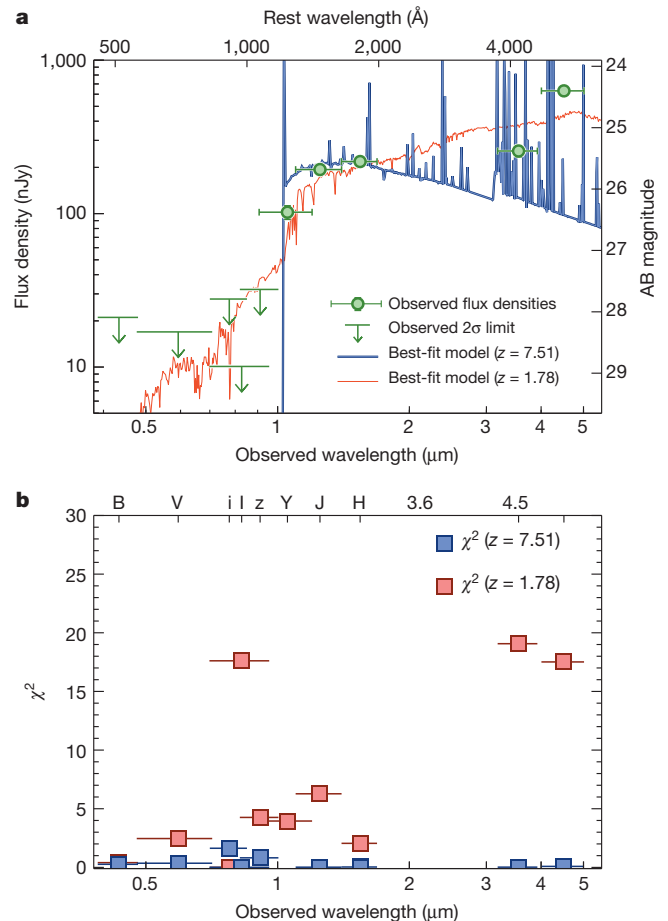


Figure 3 | Spectral energy distribution fitting of $z8_GND_5296$. **a**, The results of fitting stellar population models to the observed photometry of $z8_GND_5296$. The best-fit model for $z = 7.51$ (if the detected emission line is $\text{Ly}\alpha$) is shown by the blue spectrum, while the alternative redshift of $z = 1.78$ (if the line is $[\text{O II}]$) is shown by the red spectrum. The vertical error bars show the 1σ flux errors, while the horizontal error bars (in both panels) denote the bandpass FWHM covered by the filter. **b**, The measured χ^2 for each band for the best-fit model at each redshift. The lack of detectable optical flux, particularly in the deep F814W image, as well as the extremely red IRAC colour, strongly favour the high-redshift solution (reduced $\chi^2[z = 7.51] = 0.8$ versus $\chi^2[z = 1.78] = 14.7$). Additionally, the low-redshift model exhibits no star formation, thus this stellar population should not have detectable $[\text{O II}]$ emission. The best-fitting high-redshift model shows that this galaxy has a stellar mass of about $10^9 M_\odot$, with a 10-Myr-averaged SFR of $\sim 330 M_\odot \text{yr}^{-1}$ (68% confidence limits, $320\text{--}1,040 M_\odot \text{yr}^{-1}$). The large SFR may be responsible for the ability of $\text{Ly}\alpha$ to escape this galaxy.

enough oxygen to produce strong emission lines. Of the metallicities available in our models (0.02, 0.2, 0.4 and 1.0 times solar), only models with a metal abundance of about 20–40% of solar have $[\text{O III}]$ equivalent widths $> 300\ \text{\AA}$. Thus, even at such early times, a moderately chemically enriched galaxy could form. However, because of the discreteness of the model metallicities, further analysis is needed to draw more quantitative conclusions about the metallicity—particularly its lower limit. We note that at $z = 7.51$ $[\text{O II}]$ is in the $3.6\ \mu\text{m}$ band, but it is predicted to be about five times fainter than $[\text{O III}]$ and thus does not significantly affect the $3.6\ \mu\text{m}$ flux.

The galaxy $z8_GND_5296$ is forming stars at a very high rate, with a ‘mass-doubling’ time of at most 4 Myr. The most recent estimates¹⁷ at $z \approx 7$ find that galaxies with stellar masses of $5 \times 10^9 M_\odot$ typically have specific SFRs (that is, SFR divided by stellar mass) of $\sim 10^{-8} \text{ yr}^{-1}$. This galaxy is a factor of five less massive, yet its specific SFR is a factor of 30 greater at $3 \times 10^{-7} \text{ yr}^{-1}$, implying that $z8_GND_5296$ is undergoing a significant starburst. Additionally, estimates of the SFR functions⁹ show

Table 1 | Ly α spectroscopically confirmed galaxies at $z > 7$

ID*	$z_{\text{Ly}\alpha}$	M_{UV}^\dagger (mag)	Rest equiv. width of Ly α (Å)	SFR ‡ ($M_\odot \text{ yr}^{-1}$)	log[Stellar mass (M_\odot)]
z8_GND_5296	7.508	−21.2	8	330	9.0
SXDF-NB1006-2 (ref. 4)	7.215	−22.4 §	15	56 §	NA
GN 108036 (ref. 3)	7.213	−21.8	33	100	8.8
BDF-3299 (ref. 1)	7.109	−20.6	50	9	NA
A1703_zD6 (ref. 2)	7.045	−19.4	65	4	NA
BDF-521 (ref. 1)	7.008	−20.6	64	9	NA
IOK-1 (refs 3, 6)	6.965	−21.6	43	10	NA
HFLS3 (ref. 26)	6.337	NA	NA	2,900	10.6

NA, not available in the literature.

* Currently known galaxies with $z_{\text{Ly}\alpha} > 7$. We include IOK-1 for comparison, as it was the highest-redshift spectroscopically confirmed galaxy for several years, and HFLS3, which has the most extreme SFR known, and may represent the $z \approx 6$ evolution of z8_GND_5296.

† We compute ultraviolet absolute magnitudes (M_{UV}) for BDF-3299 and BDF-521 using the Ly α -corrected Y-band magnitudes, and for A1703_zD6 using the de-lensed J-band magnitude.

‡ The SFR for z8_GND_5296 and GN 108036 were both calculated via SED fitting. The SFR for IOK-1 was measured from Ly α emission, which is likely to be a lower limit, owing to unknown absorption. The SFRs for BDF-3299, BDF-521, A1703_zD6 and SXDF-NB1006-2 were calculated from the ultraviolet luminosity, which are also likely to be lower limits, as the ultraviolet luminosity was not corrected for dust attenuation, and the scaling relation was defined for a stellar population with an age of 100 Myr (ref. 27). The SFR for HFLS3 was derived via the infrared luminosity.

§ SXDF-NB1006-2 was only photometrically detected in a narrow band which encompassed Ly α emission. The corresponding ultraviolet absolute magnitude, and subsequent SFR, are thus highly uncertain, with published uncertainties of $M_{\text{UV}} = -22.4^{+0.4}_{-0.4}$ (ref. 4).

that a typical galaxy at $z \approx 7$ has $\text{SFR} = 10 M_\odot \text{ yr}^{-1}$; the measured SFR of z8_GND_5296 is a factor of more than 30 times greater. If this SFR function is accurate, the expected space density per co-moving Mpc^3 for this galaxy would be $\ll 10^{-5}$. The implied rarity of this galaxy could indicate that it is the progenitor of some of the most massive systems in the high-redshift Universe. However, the $z = 7.213$ galaxy GN 108036 (ref. 3), also in the GOODS-North field, also has an implied $\text{SFR} > 100 M_\odot$. Although the current statistics are poor, the presence of these two galaxies in a relatively small survey area suggests that the abundance of galaxies with such large SFRs may have previously been underestimated. If the high SFR of z8_GND_5296 continues down to $z = 6.3$, it would have a stellar mass of $\sim 5 \times 10^{10} M_\odot$, comparable to the extreme star forming $z = 6.34$ galaxy HFLS3 (Table 1)¹⁸. Should z8_GND_5296 in fact be a progenitor of such submillimetre galaxies, it is probably in the process of enshrouding itself in dust.

Both z8_GND_5296 and GN 108036 also have young inferred ages and IRAC colours indicative of strong [O III] emission. Given the difficulty of detecting Ly α emission at $z \geq 6.5$, it is interesting that these highest-redshift Ly α -detected galaxies appear to have extreme SFRs and high [O III] emission. It may be that a high SFR and/or a high excitation are necessary conditions for Ly α escape in the distant Universe—perhaps through blowing holes in the interstellar medium (ISM), allowing both Ly α and ionizing photons to escape. An outflow in the ISM of $200\text{--}300 \text{ km s}^{-1}$ could clear a hole in this galaxy in about 3–5 Myr, or perhaps even sooner if the galaxy is undergoing a merger, which could preferentially clear some lines of sight for Ly α to escape.

Finally, we examine the lack of detected Ly α lines in our full data set. If the Ly α equivalent width distribution continues its observed increase¹⁹ from $3 < z < 6$ out to $z \approx 7\text{--}8$, we should have detected Ly α emission from six galaxies. Our single detection rules out this equivalent width distribution at 2.5σ significance. This confirms previous results at $z \approx 6.5$ (refs 3, 5, 6 and 8), but here we probe $z > 7$. The lack of detectable Ly α emission is unlikely to be due to sample contamination, as contamination by lower-redshift interlopers is probably not dominant at $z = 7$ given the low contamination rate at $z = 6$ (ref. 8). To explain the low detection rate of Ly α , a neutral fraction in the intergalactic medium (IGM) at $z = 6.5$ as high as 60–90% has been proposed^{3,8}, implying a rapid increase from $z = 6$ (ref. 20). However, most other observations are consistent with an IGM neutral fraction $\leq 10\%$ at $z = 7$ (refs 21, 22), thus alternative explanations for the dearth of Ly α emission need to be explored.

One alternative explanation for at least part of the Ly α deficit may be gas within galaxies. A high ratio of gas mass to stellar mass may be consistent with the very high SFR of z8_GND_5296, as galaxies should not have SFRs (for long periods) exceeding their average gas accretion rate from the IGM (which is set by the total baryonic mass). For the inferred stellar mass and redshift, z8_GND_5296 must have a gas reservoir of

about 50 times the stellar mass to give an accretion rate comparable to the SFR²³. If true, this galaxy would have a gas surface density similar to the most gas-rich galaxies in the local Universe, and its SFR would be consistent with local relations between the gas and SFR surface densities²⁴. The large gas-to-stellar mass ratio could be due to low metallicities at earlier times which may initially inhibit star formation, allowing the formation of such a large gas reservoir²⁵. If such high gas-to-stellar mass ratios are common amongst $z > 7$ galaxies, it could explain the relative paucity of Ly α emission in our observations. Direct observations of the gas properties of distant galaxies are required to make progress in understanding both the fuelling of star formation and the escape of Ly α photons.

Received 7 June; accepted 13 September 2013.

1. Vanzella, E. *et al.* Spectroscopic confirmation of two Lyman break galaxies at redshift beyond 7. *Astrophys. J.* **730**, L35 (2011).
2. Schenker, M. A. *et al.* Keck spectroscopy of faint $3 < z < 8$ Lyman break galaxies: evidence for a declining fraction of emission line sources in the redshift range $6 < z < 8$. *Astrophys. J.* **744**, 179 (2012).
3. Ono, Y. *et al.* Spectroscopic confirmation of three z-dropout galaxies at $z = 6.844\text{--}7.213$: demographics of Ly α emission in $z \sim 7$ galaxies. *Astrophys. J.* **744**, 83 (2012).
4. Shibuya, T. *et al.* The first systematic survey for Ly α emitters at $z = 7.3$ with red-sensitive Subaru/Suprime-Cam. *Astrophys. J.* **752**, 114 (2012).
5. Kashikawa, N. *et al.* The end of the reionization epoch probed by Ly α emitters at $z = 6.5$ in the Subaru deep field. *Astrophys. J.* **648**, 7–22 (2006).
6. Iye, M. *et al.* A galaxy at a redshift $z = 6.96$. *Nature* **443**, 186–188 (2006).
7. Ouchi, M. *et al.* Statistics of 207 Ly α emitters at a redshift near 7: constraints on reionization and galaxy formation models. *Astrophys. J.* **723**, 869–894 (2010).
8. Pentericci, L. *et al.* Spectroscopic confirmation of $z \sim 7$ Lyman break galaxies: probing the earliest galaxies and the epoch of reionization. *Astrophys. J.* **743**, 132 (2011).
9. Smit, R. *et al.* The star formation rate function for redshift $z \sim 4\text{--}7$ galaxies: evidence for a uniform buildup of star-forming galaxies during the first 3 Gyr of cosmic time. *Astrophys. J.* **756**, 14 (2012).
10. Grogin, N. A. *et al.* CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. *Astrophys. J.* **197** (Supp.), 35 (2011).
11. Koekemoer, A. M. *et al.* CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey – The Hubble Space Telescope observations, imaging data products, and mosaics. *Astrophys. J.* **197** (Supp.), 36 (2011).
12. McLean, I. S. *et al.* MOSFIRE, the multi-object spectrometer for infra-red exploration at the Keck Observatory. *Proc. SPIE* **8446**, 84460–1–15 (2012).
13. Finkelstein, S. L. *et al.* On the stellar populations and evolution of star-forming galaxies at $6.3 \leq z \leq 8.6$. *Astrophys. J.* **719**, 1250–1273 (2010).
14. McLure, R. J. *et al.* Galaxies at $z = 6\text{--}9$ from the WFC3/IR imaging of the Hubble Ultra Deep Field. *Mon. Not. R. Astron. Soc.* **403**, 960–983 (2010).
15. Finkelstein, S. L. *et al.* CANDELS: The evolution of galaxy rest-frame ultraviolet colors from $z = 8$ to 4. *Astrophys. J.* **756**, 164 (2012).
16. McLure, R. J. *et al.* A new multifield determination of the galaxy luminosity function at $z = 7\text{--}9$ incorporating the 2012 Hubble Ultra-Deep Field imaging. *Mon. Not. R. Astron. Soc.* **432**, 2696–2716 (2013).
17. Stark, D. P. *et al.* Keck spectroscopy of $3 < z < 7$ faint Lyman break galaxies: the importance of nebular emission in understanding the specific star formation rate and stellar mass density. *Astrophys. J.* **763**, 129 (2013).
18. Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).

19. Stark, D. P., Ellis, R. S. & Ouchi, M. Keck spectroscopy of faint $3 < z < 7$ Lyman break galaxies: a high fraction of line emitters at redshift six. *Astrophys. J.* **728**, L2 (2011).
20. Fan, X. *et al.* Constraining the evolution of the ionizing background and the epoch of reionization with $z \sim 6$ quasars. II. A sample of 19 quasars. *Astron. J.* **132**, 117–136 (2006).
21. Bolton, J. S. *et al.* How neutral is the intergalactic medium surrounding the redshift $z = 7.085$ quasar ULASJ1120+0641? *Mon. Not. R. Astron. Soc.* **416**, L70–L74 (2011).
22. Finkelstein, S. L. *et al.* CANDELS: the contribution of the observed galaxy population to cosmic reionization. *Astrophys. J.* **758**, 93 (2012).
23. Dekel, A. *et al.* Toy models for galaxy formation versus simulations. *Mon. Not. R. Astron. Soc.* **435**, 999–1019 (2013).
24. Kennicutt, R. C. & Evans, N. J. Star formation in the Milky Way and nearby galaxies. *Annu. Rev. Astron. Astrophys.* **50**, 531–608 (2012).
25. Krumholz, M. R. & Dekel, A. Metallicity-dependent quenching of star formation at high redshift in small galaxies. *Astrophys. J.* **753**, 16 (2012).
26. Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).
27. Kennicutt, R. C. Jr. Star formation in galaxies along the Hubble sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–231 (1998).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Dijkstra, J. Rhoads and S. Malhotra for conversations, as well as N. Konidaris and C. Steidel for assistance with the MOSFIRE data reduction pipeline. We also thank our Keck Support Astronomer G. Wirth for assistance during our observing run. S.L.F. acknowledges support from the University of Texas at Austin, the McDonald Observatory and NASA through a NASA Keck PI Data Award, administered

by the NASA Exoplanet Science Institute. Data presented here were obtained at the W. M. Keck Observatory from telescope time allocated to NASA through the agency's scientific partnership with the California Institute of Technology and the University of California. The Observatory was made possible by the financial support of the W. M. Keck Foundation. We recognize and acknowledge the cultural role and reverence that the summit of Mauna Kea has within the indigenous Hawaiian community. This work is also based in part on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555, as well as the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA.

Author Contributions S.L.F. wrote the text, obtained and reduced the data and led the initial observing proposal. C.P. and M.D. assisted with the analysis of the data. M.S. and V.T. assisted with the observation planning and implementation. K.D.F. performed the Spitzer/IRAC photometry. A.M.K. was responsible for the reduction of the optical and NIR imaging data used to select the sample. G.G.F., M.L.N.A. and S.P.W. obtained and reduced the mid-infrared data. B.J.W. provided grism spectroscopic information. B.M., H.C.F., M.G., N.R., A.D., A.F., N.A.G., J.-S.H., D.K. and M.R. have contributed in their roles as members of the CANDELS and S-CANDELS teams, and assisted with the planning and interpretation of the observations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.L.F. (stevenf@astro.as.utexas.edu).

Interface superconductor with gap behaviour like a high-temperature superconductor

C. Richter^{1,2}, H. Boschker¹, W. Dietsche¹, E. Fillis-Tsirakis¹, R. Jany², F. Loder², L. F. Kourkoutis^{3,4}, D. A. Muller^{3,4}, J. R. Kirtley⁵, C. W. Schneider⁶ & J. Mannhart¹

The physics of the superconducting state in two-dimensional (2D) electron systems is relevant to understanding the high-transition-temperature copper oxide superconductors and for the development of future superconductors based on interface electron systems¹. But it is not yet understood how fundamental superconducting parameters, such as the spectral density of states, change when these superconducting electron systems are depleted of charge carriers. Here we use tunnel spectroscopy with planar junctions to measure the behaviour of the electronic spectral density of states as a function of carrier density, clarifying this issue experimentally. We chose the conducting LaAlO₃–SrTiO₃ interface² as the 2D superconductor, because this electron system can be tuned continuously with an electric gate field³. We observed an energy gap of the order of 40 micro-electronvolts in the density of states, whose shape is well described by the Bardeen–Cooper–Schrieffer superconducting gap function. In contrast to the dome-shaped dependence of the critical temperature, the gap increases with charge carrier depletion in both the underdoped region and the overdoped region. These results are analogous to the pseudogap behaviour of the high-transition-temperature copper oxide superconductors and imply that the smooth continuation of the superconducting gap into pseudogap-like behaviour could be a general property of 2D superconductivity.

One of the main challenges in understanding the superconductivity of copper oxide superconductors is to identify the origin and nature of the pseudogap phase in the underdoped regime^{4,5}. This phase is characterized by a reduction in the density of states (DOS) at the Fermi level on a part of the Fermi surface⁶. With decreasing doping, this reduction increases and persists up to temperatures well above the resistive critical temperature, T_c . The pseudogap is linked to the correlations in the electron systems emerging from the doping of the antiferromagnetically ordered Mott state of the undoped parent compounds⁷. Such correlated electron systems are susceptible to various electronic instabilities, such as static stripe formation⁸, quantum liquid-crystal order^{9,10}, spontaneous diamagnetism¹¹ and incommensurate charge fluctuations¹². The pseudogap behaviour has been attributed¹³ to these phases, and their competition with superconductivity is thought to result in the reduction in T_c . However, it has also been suggested that the pseudogap results from preformed Cooper pairs and thus has its origin in superconductivity^{14–16}. This would imply that Cooper pairs are present in the pseudogap phase, yet the pairs lack the phase coherence required for macroscopic superconductivity¹⁷. Here we report that the doping dependence of the high- T_c -superconductor pseudogap is shared by the electron liquid at the LaAlO₃–SrTiO₃ interface, which is a model 2D superconductor that is much simpler than the copper oxide superconductors: it is not, for example, a doped Mott insulator.

Normal metal–insulator–superconductor tunnel spectroscopy provides direct experimental access to the superconducting gap because the differential conductance characteristic, $dI/dV(V)$, of a tunnel junction is a measure of the spectral DOS of the electrons in the superconductor

close to the Fermi energy, E_F (ref. 18). Here, I is the tunnel current and V is the voltage between the superconductor and the metallic counter electrode. To study the evolution of the superconducting gap, we developed a tunnel structure in which the LaAlO₃ is used to generate the two-dimensional electron liquid (2DEL) and simultaneously to act as tunnel barrier. Figure 1a and Fig. 1b respectively show a photograph and a cross-sectional sketch of the tunnel junctions. The LaAlO₃ layers (5.6-eV bandgap) were deposited on TiO₂-terminated SrTiO₃ substrates. To obtain a sizable tunnel current, the LaAlO₃ films were usually grown to a thickness of four monolayers of LaAlO₃, the minimum thickness required to induce the 2DEL. On top of the LaAlO₃ film, we deposited metallic layers of Au (Methods). To minimize the density of impurity interface states, all deposition processes were performed *in situ*. To be able to perform four-point measurements of the tunnel characteristics, the samples were patterned into the circular geometry shown in Fig. 1. The top electrode is provided by a Au ring with an inner diameter of 160 μm and an outer diameter of 400–1,000 μm . Two ohmic contacts connect to the 2DEL; one contact is made by a disk inside the top-electrode ring and the other is made by an outer ring. Gate fields are generated by voltages, V_G , applied to the back of the SrTiO₃ substrates. A total of seven devices were studied and showed consistent results.

The microstructure of the tunnel barrier was assessed by scanning transmission electron microscopy (STEM) in combination with electron energy-loss spectroscopy. The high-angle annular dark-field STEM image (Fig. 1c) and the chemical map (Extended Data Fig. 1) show the LaAlO₃ tunnel barrier to be homogeneous; no pinholes penetrating the barrier are seen. Although the LaAlO₃ layer at the interface to the Au electrode is disordered, the microstructure at the LaAlO₃–SrTiO₃ interface is well preserved with most of the interdiffusion confined to within one unit cell on either side of the interface. Diffusion of Au towards the interface is not observed. At 4.2 K, we characterized the tunnel junctions in the normal state in the voltage range from –200 to 100 mV (Fig. 2a, b). We observed asymmetric tunnel characteristics with a large tunnel current for $V > 0$ and a smaller tunnel current for $V < 0$. The polarity of the voltage reflects the sign of the interface voltage with respect to the top electrode bias. For $V < 0$, electrons tunnel out of the 2DEL and the $dI/dV(V)$ characteristic is shaped by inelastic tunnelling processes. The data show the tunnelling to be assisted by SrTiO₃ longitudinal optical phonons¹⁹, which generate peaks in the $dI/dV(V)$ characteristics at energies corresponding to large phonon densities of states: ~ 60 and ~ 100 meV. The presence of phonon peaks in the tunnel characteristics provides evidence that the electron transport across the LaAlO₃ happens by means of tunnelling, where for $V > 0$ elastic tunnelling predominates. The $dI/dV(V)$ characteristics show a strong energy dependence around E_F , which lies only ~ 20 meV above the minimum of the characteristics. We interpret this minimum to mark the edge of the conduction band of the LaAlO₃–SrTiO₃ interface 2DEL.

¹Max Planck Institute for Solid State Research, 70569 Stuttgart, Germany. ²Experimental Physics VI, Center for Electronic Correlations and Magnetism, Augsburg University, 86135 Augsburg, Germany.

³School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA. ⁴Kavli Institute at Cornell for Nanoscale Science, Ithaca, New York 14853, USA. ⁵Center for Probing the Nanoscale, Stanford University, Stanford, California 94305-4045, USA. ⁶Paul Scherrer Institut, 5232 Villigen, Switzerland.

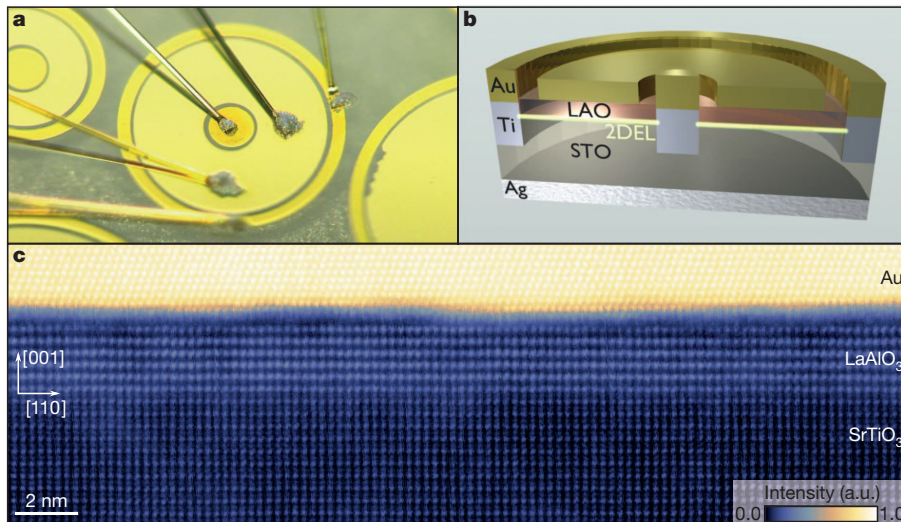


Figure 1 | Device layout. **a**, **b**, Photograph (**a**) and schematic cross section (**b**) of a typical Au–LaAlO₃–SrTiO₃ tunnel device. The broad gold ring (inner diameter, 160 μm) lies on top of the LaAlO₃ layer, which serves as a tunnel barrier between the 2DEL and the Au. The outer ring and the centre contact of

the device are Au-covered Ti contacts to the 2DEL. **c**, Cross-sectional high-angle annular dark-field STEM image of a Au–LaAlO₃–SrTiO₃ tunnel junction. The image is taken along the [110] zone axis of the perovskite unit cells. a.u., arbitrary units.

At temperatures below T_c , the tunnelling current provides direct information on the superconducting state. Figure 2c shows typical $dI/dV(V)$ tunnelling characteristics measured below 100 μV . The characteristics reveal a clear superconducting gap, Δ , for which an analysis using a Dynes fit²⁰ with a single s -wave gap as fitting parameter yields $\Delta(0\text{ K}) = (40 \pm 2)\text{ }\mu\text{eV}$ (Extended Data Fig. 2). The gap and the coherence peaks are well developed and are consistent with a laterally homogeneous superconducting state. The gap closes at $T_{\text{gap}} = 0.28\text{ K}$. A detailed analysis of the temperature dependence of the spectra is given in Methods. We note that our observation of the superconducting gap implies that no conducting layer exists between the superconducting sheet and the LaAlO₃ layers. We did not observe signatures of a second superconducting gap as reported²¹ for superconducting Nb-doped SrTiO₃.

Having stated these results, we now address the main question of our study, namely how the superconducting gap relates to the superconducting transition temperature in a 2D superconductor tuned by electrostatic field effect doping. To compare the size of the superconducting gap directly with the transition temperature, we fabricated

devices with four contacts to the 2DEL as well, enabling measurements of the 2DEL resistance within the tunnel device (Methods). We found that positive gate voltages (carrier accumulation) enhance the DOS at E_F and suppress the coherence peaks. Negative gate voltages (carrier depletion) suppress the DOS at E_F and broaden the coherence peaks (Fig. 3a). Moreover, the coherence peak maxima shift systematically to higher voltages with decreasing carrier density, indicating an increase of the superconducting gap. The temperature, T_{gap} , at which the gap closes increases with decreasing charge carrier density (Fig. 3b and Extended Data Fig. 3), approximately following the low-temperature value of the gap. The temperature dependence of Δ is BCS-like with $2\Delta/k_B T_{\text{gap}} \approx 3.4$. Here k_B is Boltzmann's constant.

The gate voltage dependence of Δ , T_{gap} and T_c is presented in Fig. 4a, b. The transition temperature does not follow Δ and T_{gap} ; Δ and T_{gap} increase with charge carrier depletion over the entire voltage range, whereas T_c has a dome-shaped dependence. A maximum T_c of 0.27 K is observed at $V_G = 0\text{ V}$, for $V_G > 0\text{ V}$ the 2DEL is overdoped and for $V_G < 0\text{ V}$ the 2DEL is underdoped. For $V_G < -150\text{ V}$, we did not observe a superconducting transition in the temperature-dependent

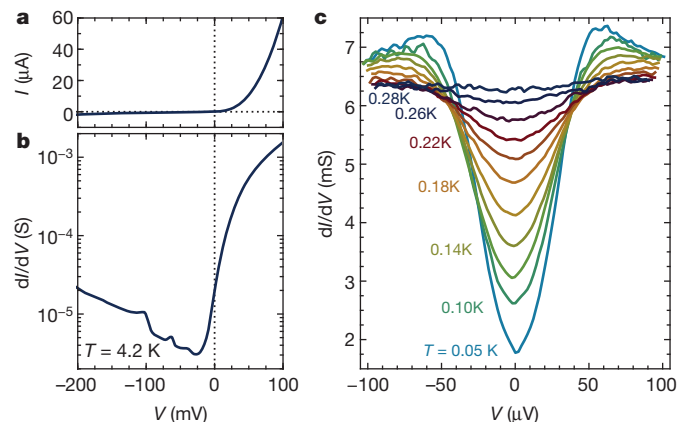


Figure 2 | Large-range tunnel spectra and the superconducting gap. **a**, Current-versus-voltage tunnel characteristic, $I(V)$, measured at 4.2 K. The voltage characterizes the voltage applied to the interface 2DEL; a positive current is provided by electrons tunnelling from the Au into the 2DEL. **b**, Differential conductance, $dI/dV(V)$, measured in the normal-conducting state (4.2 K). **c**, Temperature-dependent tunnel spectra in the superconducting state. The gap closes at 0.28 K. The device area is 0.3 mm^2 .

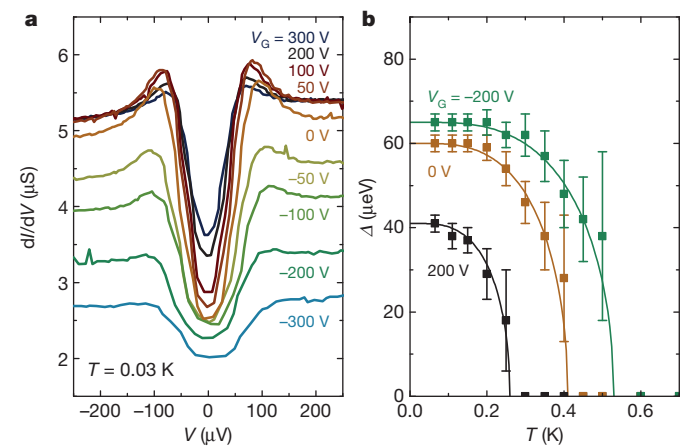


Figure 3 | Dependence of the tunnel spectra on gate voltage. **a**, Tunnel spectra as a function of the back-gate voltage, V_G (positive voltage corresponds to carrier accumulation). The device area is 0.5 mm^2 . **b**, Temperature dependence of Δ for different values of V_G . The solid lines are the predictions of the BCS model. Error bars define the 90% confidence interval.

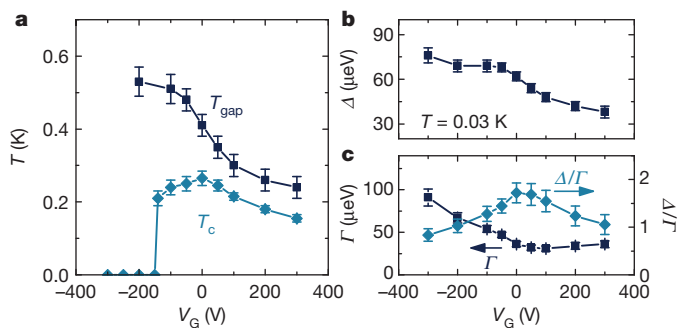


Figure 4 | Dependences of T_c and T_{gap} on gate voltage. Measured dependence on gate voltage of the superconducting transition temperature, T_c , and the temperature at which the gap closes, T_{gap} (a); the superconducting gap, Δ (0 K) (b); and the coherence-peak-broadening parameter, Γ (0 K), and the ratio Δ (0 K)/ Γ (0 K) (c). Error bars define the 90% confidence interval.

resistance, $R(T)$, measurements (Methods and Extended Data Fig. 4). Because the gap is still observed in the tunnel characteristics for $V_G < -150$ V, it is either present in the insulating state, for example as expected from the observation of the superconductor–insulator transition at the quantum pair resistance³; or the superconductivity in the device is inhomogeneous, with normal-conducting or insulating areas and areas that are still superconducting. Inhomogeneous transport in the normal state was recently observed in the 2DEL. However, the inhomogeneity of the superfluid density was reported to be small and to diminish with increasing negative gate voltage²². Figure 4c shows the coherence-peak-broadening parameter²⁰, Γ , obtained from the Dynes fit, and the ratio Δ/Γ , which can be interpreted as the product of the strength of the superconducting pairing interaction and the strength of the quasiparticle coherence. As a function of the charge carrier density, the critical temperature follows the Δ/Γ ratio well.

Contrary to our expectations, Δ and T_{gap} do not have a maximum value that coincides with the maximum T_c of the superconducting dome, but rather the gap increases continuously with charge carrier depletion. Because the gap continuously evolves from the overdoped region to the underdoped region and the quasiparticle peaks remain present, we conclude that the gap in the underdoped regime is indeed a superconducting gap. The reduction in T_c with respect to T_{gap} can then be due either to a competing order parameter or to weak phase coherence in the superconductor. The observation that T_c scales with Δ/Γ indicates that the limited quasiparticle lifetime is an important factor controlling T_c in the underdoped regime. In the case of a superconductor limited by phase coherence, T_c is expected to be proportional to the superfluid density²³. Recent measurements²⁴ demonstrated that the superfluid density decreases with charge carrier depletion in the underdoped regime, indicating the importance of phase fluctuations in this part of the phase diagram. In addition, the reduction in T_c can also be explained by a competing order parameter. This phase is then expected to involve a mechanism that enables an additional quasiparticle scattering channel, thereby reducing its lifetime.

Figure 5 illustrates our main result: the doping-dependent pseudogap behaviour of the high- T_c copper oxide superconductors is analogous to the gap behaviour of the LaAlO_3 – SrTiO_3 interface 2DEL, even though the superconducting dome of the latter system occurs for a carrier density ten times lower than that of the former system^{3,25}. In both systems, T_{gap} does not follow T_c in the underdoped region of the phase diagram, but increases with charge carrier depletion. Moreover a reduction in the quasiparticle lifetime has been observed in the underdoped region of the high- T_c copper oxide superconductors^{6,26}, very similar to our result for the non-copper-oxide interface 2DEL. These commonalities show that much of the high- T_c -superconductor pseudogap behaviour is found in a 2D superconductor that has a completely different Fermi surface²⁷ than the high- T_c superconductors and in which, in contrast to the high- T_c superconductors, there are no Mott

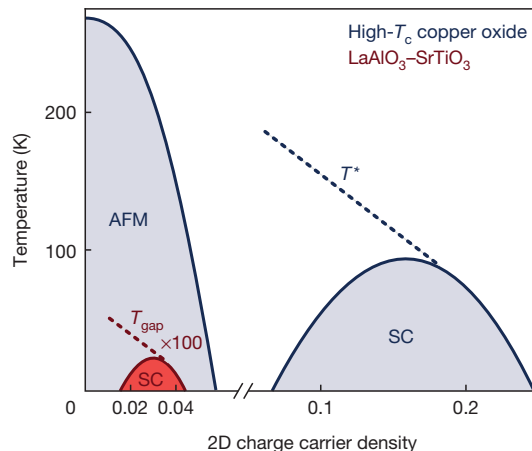


Figure 5 | Comparison between phase diagrams for LaAlO_3 – SrTiO_3 and copper oxide superconductors. Illustration of the doping-versus-temperature phase diagram of the n-doped LaAlO_3 – SrTiO_3 interface 2DEL and the p-doped high- T_c copper oxide superconductors. The charge carrier density is given in units of charge carriers per 2D unit cell. AFM, antiferromagnetic; SC, superconducting.

phases and antiferromagnetic insulating states in the underdoped regime. The common presence of a gap above T_c in 2D systems, for example in ultrathin TiN films²⁸ and in atomic Fermi gases²⁹, and the behaviour of the LaAlO_3 – SrTiO_3 interface superconductor therefore suggest that 2D superconductors in general have a phase diagram with a gap above T_c in the underdoped region and an ever increasing gap with charge carrier depletion.

METHODS SUMMARY

Using pulsed laser deposition monitored by reflection high-energy electron diffraction, LaAlO_3 films were grown onto TiO_2 -terminated SrTiO_3 substrates (CrysTec GmbH) at an oxygen pressure of 1×10^{-4} mbar at 780 °C. The LaAlO_3 was ablated from a single-crystalline target with a laser fluence of $\sim 1 \text{ J cm}^{-2}$. After annealing, the samples were transferred *in situ* into a sputtering system, where ~ 30 nm of Au were deposited onto the sample surface by radio-frequency sputtering. The Au was patterned subsequently into photolithographically defined, ring-shaped electrodes by wet etching with a $\text{KI} + \text{I}_2$ solution. Contacts to the 2DEL were made by refilling Ar-ion-etched pits with sputtered Ti and Au. We attached wires to the top electrode using Ag glue and to the contacts to the 2DEL with wedge bonding. Tunnelling spectra were acquired by sourcing current from the top Au electrode to the centre contact of the 2DEL and measuring the voltage between a second wire on the top contact and the outer 2DEL contact ring. To tune the superconducting state electrostatically, a gate voltage, V_G , was applied to the Ag-coated back side of the SrTiO_3 substrate while the 2DEL was held at ground potential. The electron microscopy and spectroscopy measurements were performed on the aberration-corrected 100-kV Nion UltraSTEM at Cornell University (Extended Data Fig. 1). Each spectrum in the 180×180 pixel map was acquired for 20 ms and captured the $\text{Ti L}_{2,3}$ edge, the O K edge and the $\text{La M}_{4,5}$ edge simultaneously. One of the samples was grown on a SrTiO_3 substrate in which the oxygen ions were partly exchanged for the heavier ^{18}O isotope before film growth³⁰. No clear difference in the superconducting properties was found between this sample and the other samples, either in tunnelling or in $R(T)$ measurements.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 March; accepted 15 July 2013.

Published online 6 October 2013.

- Pereiro, J., Petrovic, A., Panagopoulos, C. & Božović, I. Interface superconductivity: history, development and prospects. *Phys. Express* **1**, 208–241 (2011).
- Ohtomo, A. & Hwang, H. Y. A high-mobility electron gas at the LaAlO_3 / SrTiO_3 heterointerface. *Nature* **427**, 423–426 (2004).
- Caviglia, A. D. *et al.* Electric field control of the LaAlO_3 / SrTiO_3 interface ground state. *Nature* **456**, 624–627 (2008).

4. Warren, W. W. *et al.* Cu spin dynamics and superconducting precursor effects in planes above T_c in $\text{YBa}_2\text{Cu}_3\text{O}_{6.7}$. *Phys. Rev. Lett.* **62**, 1193–1196 (1989).
5. Ding, H. *et al.* Spectroscopic evidence for a pseudogap in the normal state of underdoped high- T_c superconductors. *Nature* **382**, 51–54 (1996).
6. Damascelli, A., Hussain, Z. & Shen, Z.-X. Angle-resolved photoemission studies of the cuprate superconductors. *Rev. Mod. Phys.* **75**, 473–541 (2003).
7. Lee, P. A., Nagaosa, N. & Wen, X.-G. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78**, 17–85 (2006).
8. Tranquada, J. M., Sternlieb, B. J., Axe, J. D., Nakamura, Y. & Uchida, S. Evidence for stripe correlations of spins and holes in copper oxide superconductors. *Nature* **375**, 561–563 (1995).
9. Kivelson, S. A., Fradkin, E. & Emery, V. J. Electronic liquid-crystal phases of a doped Mott insulator. *Nature* **393**, 550–553 (1998).
10. Hinkov, V. *et al.* Electronic liquid crystal state in the high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6.45}$. *Science* **319**, 597–600 (2008).
11. Fauqué, B. *et al.* Magnetic order in the pseudogap phase of high- T_c superconductors. *Phys. Rev. Lett.* **96**, 197001 (2006).
12. Ghiringhelli, G. *et al.* Long-range incommensurate charge fluctuations in $(\text{Y,Nd})\text{Ba}_2\text{Cu}_3\text{O}_{6+x}$. *Science* **337**, 821–825 (2012).
13. Norman, M. R., Pines, D. & Kallin, C. The pseudogap: friend or foe of high- T_c ? *Adv. Phys.* **54**, 715–733 (2005).
14. Orenstein, J. & Millis, A. J. Advances in the physics of high temperature superconductors. *Science* **288**, 468–474 (2000).
15. Renner, Ch., Revaz, B., Genoud, J.-Y., Kadowaki, K. & Fisher, Ø. Pseudogap precursor of the superconducting gap in under- and overdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Phys. Rev. Lett.* **80**, 149–152 (1998).
16. Miyakawa, N. *et al.* Predominantly superconducting origin of large energy gaps in underdoped $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ from tunneling spectroscopy. *Phys. Rev. Lett.* **83**, 1018–1021 (1999).
17. Emery, V. J. & Kivelson, S. A. Importance of phase fluctuations in superconductors with small superfluid density. *Nature* **374**, 434–437 (1995).
18. Giaever, I. Energy gap in superconductors measured by electron tunneling. *Phys. Rev. Lett.* **5**, 147–148 (1960).
19. Servoin, J. L., Luspain, Y. & Gervais, F. Infrared dispersion in SrTiO_3 at high temperature. *Phys. Rev. B* **22**, 5501–5506 (1980).
20. Dynes, R. C., Narayanamurti, V. & Garno, J. P. Direct measurement of quasiparticle-lifetime broadening in a strong-coupled superconductor. *Phys. Rev. Lett.* **41**, 1509–1512 (1978).
21. Binnig, G., Baratoff, A., Hoenig, H. E. & Bednorz, J. G. Two-band superconductivity in Nb-doped SrTiO_3 . *Phys. Rev. Lett.* **45**, 1352–1355 (1980).
22. Kalisky, B. *et al.* Locally enhanced conductivity due to tetragonal domain structure in $\text{LaAlO}_3/\text{SrTiO}_3$ heterointerfaces. *Nature Mater.* (in the press).
23. Uemura, Y. J. *et al.* Universal correlations between T_c and n_s/m^* (carrier density over effective mass) in high- T_c cuprate superconductors. *Phys. Rev. Lett.* **62**, 2317–2320 (1989).
24. Bert, J. A. *et al.* Gate-tuned superfluid density at the superconducting $\text{LaAlO}_3/\text{SrTiO}_3$ interface. *Phys. Rev. B* **86**, 060503(R) (2012).
25. Fischer, Ø., Kugler, M., Maggio-Aprile, I., Berthod, C. & Renner, Ch. Scanning tunneling spectroscopy of high-temperature superconductors. *Rev. Mod. Phys.* **79**, 353–419 (2007).
26. Joshua, A., Pecker, S., Ruhman, J., Altman, E. & Ilani, S. A universal critical density underlying the physics of electrons at the $\text{LaAlO}_3/\text{SrTiO}_3$ interface. *Nature Commun.* **3**, 1129 (2012).
27. Berner, G. *et al.* Direct k -space mapping of the electronic structure in an oxide-oxide interface. *Phys. Rev. Lett.* **110**, 247601–247605 (2013).
28. Sacépé, B. *et al.* Pseudogap in a thin film of a conventional superconductor. *Nature Commun.* **1**, 140 (2010).
29. Feld, M., Fröhlich, B., Vogt, E., Koschorreck, M. & Köhl, M. Observation of a pairing pseudogap in a two-dimensional Fermi gas. *Nature* **480**, 75–78 (2011).
30. Schneider, C. W. *et al.* The origin of oxygen in oxide thin films: role of the substrate. *Appl. Phys. Lett.* **97**, 192107 (2010).

Acknowledgements We acknowledge discussions with M. Beasley, A. P. Kampf and T. Kopp, technical support from M. Hagel, I. Hagel, M. Schmid and D. Zhang, and financial support from the German Science Foundation (TRR 80). Electron microscopy and spectroscopy was performed at the Cornell Center for Materials Research (CCMR), which is an NSF MRSEC supported by NSF grant DMR-1120296.

Author Contributions C.R., R.J. and C.W.S. prepared the samples. C.R., H.B., W.D. and E.F.-S. performed the transport measurements. L.F.K. and D.A.M. performed the electron microscopy. F.L. and J.R.K. provided theoretical input for the analysis and the modelling of the tunnel characteristics. J.M. supervised the research. C.R., H.B. and J.M. wrote the manuscript. All authors contributed to the discussion and provided feedback on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M. (j.mannhart@fkf.mpg.de).

Exploiting dimensionality and defect mitigation to create tunable microwave dielectrics

Che-Hui Lee^{1,2*}, Nathan D. Orloff^{3,4*}, Turan Birol^{5*}, Ye Zhu⁵, Veronica Goian⁶, Eduard Roca⁷, Ryan Haislmaier², Eftihia Vlahos², Julia A. Mundy⁵, Lena F. Kourkoutis^{5,8}, Yuefeng Nie¹, Michael D. Biegalski⁹, Jingshu Zhang¹, Margitta Bernhagen¹⁰, Nicole A. Benedek¹¹, Yongsam Kim⁵, Joel D. Brock⁵, Reinhard Uecker¹⁰, X. X. Xi¹², Venkatraman Gopalan², Dmitry Nuzhnyy⁶, Stanislav Kamba⁶, David A. Muller^{5,8}, Ichiro Takeuchi¹³, James C. Booth³, Craig J. Fennie⁵ & Darrell G. Schlom^{1,8}

The miniaturization and integration of frequency-agile microwave circuits—relevant to electronically tunable filters, antennas, resonators and phase shifters—with microelectronics offers tantalizing device possibilities, yet requires thin films whose dielectric constant at gigahertz frequencies can be tuned by applying a quasi-static electric field¹. Appropriate systems such as $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ have a paraelectric–ferroelectric transition just below ambient temperature, providing high tunability^{1–3}. Unfortunately, such films suffer significant losses arising from defects. Recognizing that progress is stymied by dielectric loss, we start with a system with exceptionally low loss— $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases^{4,5}—in which $(\text{SrO})_2$ crystallographic shear^{6,7} planes provide an alternative to the formation of point defects for accommodating non-stoichiometry^{8,9}. Here we report the experimental realization of a highly tunable ground state arising from the emergence of a local ferroelectric instability¹⁰ in biaxially strained $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases with $n \geq 3$ at frequencies up to 125 GHz. In contrast to traditional methods of modifying ferroelectrics—doping^{1–3,11,12} or strain^{13–16}—in this unique system an increase in the separation between the $(\text{SrO})_2$ planes, which can be achieved by changing n , bolsters the local ferroelectric instability. This new control parameter, n , can be exploited to achieve a figure of merit at room temperature that rivals all known tunable microwave dielectrics³.

Ferroelectric thin films possessing a nonlinear dielectric response to a quasi-static electric field have been widely pursued for tunable dielectric devices^{17–20} that work at gigahertz frequencies. $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ is the most common of such materials because of its high tunability ($\Delta K/K$, where K is the dielectric constant and ΔK is the change in dielectric constant under the application of a quasi-static electric field) and composition-dependent Curie temperature, T_C (refs 1–3, 11, 12). Thin films, however, of $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ suffer significant dielectric losses at application-relevant operating frequencies^{1,12,21}. These losses are believed to arise from the motion of charged defects in a time-dependent electromagnetic field, as well as local polar nanoregions induced by structural imperfections and non-stoichiometry¹. These losses are significantly higher in $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ films than in the bulk material^{1,21}.

Our approach to this problem is to take a system with low loss related to $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$, and engineer it to improve its tunability. We selected the $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ Ruddlesden–Popper series of phases^{22,23}, which are known to have low loss in bulk^{4,5}. In this homologous series, the positive integer n corresponds to the number of perovskite SrTiO_3 layers that are sandwiched between double SrO rock-salt layers (Fig. 1a). Although bulk $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases are centrosymmetric^{22,23} and are thus non-polar, calculations from first principles recently predicted

that under biaxial tensile strain $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases can exhibit a ferroelectric instability²⁴, local to the perovskite layers, if the spacing between the $(\text{SrO})_2$ planes is sufficiently large (high n) to exceed a coherence length that depends on epitaxial strain¹⁰.

In Fig. 1b (right-hand axis), the square of the in-plane polar phonon frequencies of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases, calculated from first principles, are plotted as a function of n for $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ commensurately strained to the in-plane lattice parameter of a (110) DyScO_3 substrate. (Details of the calculations are provided in ref. 10 and in Methods.) It is seen that there is a critical n_c at which the square of the polar soft-mode frequency becomes negative, indicating a ferroelectric instability, and ferroelectricity for $n \geq n_c$ is expected. This n_c can also be seen from the energy lowering provided by the local ferroelectric instability (Fig. 1b, left-hand axis) and the curves of energy against total in-plane polar displacement in the $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ (Fig. 1c, d). Our calculations show that for $n \geq 3$ a local ferroelectric instability is expected at $T = 0$ for $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ commensurate with (110) DyScO_3 . (The instability at $n = 2$ is too weak to stabilize a spontaneous polarization when quantum fluctuations of nuclei are considered¹⁰.) As Fig. 1b–d shows, the in-plane ferroelectric instability of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases can be tuned by changing the out-of-plane distance between the $(\text{SrO})_2$ layers; that is, by changing the value of n . Such a control parameter is a new¹⁰ and potentially disorder-free way of manipulating the properties of a tunable dielectric.

Recognizing that the presence of $(\text{SrO})_2$ crystallographic shear planes^{6,7} in $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases with finite n could provide a means of locally accommodating non-stoichiometry^{8,9} (as described below), we investigate the tunable dielectric figure of merit²⁵ (FOM)

$$\frac{K(V=0) - K(V)}{K(V=0)\tan\delta}$$

of commensurate $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ films grown on (110) DyScO_3 with finite $n > n_c$. Although the $n = 1, 2, 3$ and $n = \infty$ members of the $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ series are the only compositions that can be synthesized in single-phase form in bulk material^{9,22,23}, by supplying incident species in an ordered sequence with submonolayer composition control, oxide molecular-beam epitaxy has enabled the growth of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ films with n as high as 10 in single-phase form²⁶, even though the formation energies of high- n phases are essentially degenerate^{27,28}.

In this study we grew epitaxial $n = 1$ to $n = 6$ $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ films on (110) DyScO_3 and (110) GdScO_3 substrates²⁹. (Details of the thin-film growth are given in Methods.) X-ray diffraction (XRD) scans (Fig. 2a and Extended Data Figs 2–4) of the $n = 1$ to $n = 6$ films reveal them to

¹Department of Materials Science and Engineering, Cornell University, Ithaca, New York 14853, USA. ²Department of Materials Science and Engineering, Pennsylvania State University, University Park, Pennsylvania 16802, USA. ³National Institute of Standards and Technology, Boulder, Colorado 80305, USA. ⁴Department of Physics, University of Maryland, College Park, Maryland 20742, USA. ⁵School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA. ⁶Institute of Physics ASCR, Na Slovance 2, 182 21 Prague 8, Czech Republic. ⁷Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain. ⁸Kavli Institute at Cornell for Nanoscale Science, Ithaca, New York 14853, USA. ⁹Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA. ¹⁰Leibniz Institute for Crystal Growth, Max-Born-Strasse 2, D-12489 Berlin, Germany. ¹¹Materials Science and Engineering Program, The University of Texas at Austin, Austin, Texas 78712, USA. ¹²Department of Physics, Temple University, Philadelphia, Pennsylvania 19122, USA. ¹³Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, USA.

*These authors contributed equally to this work.

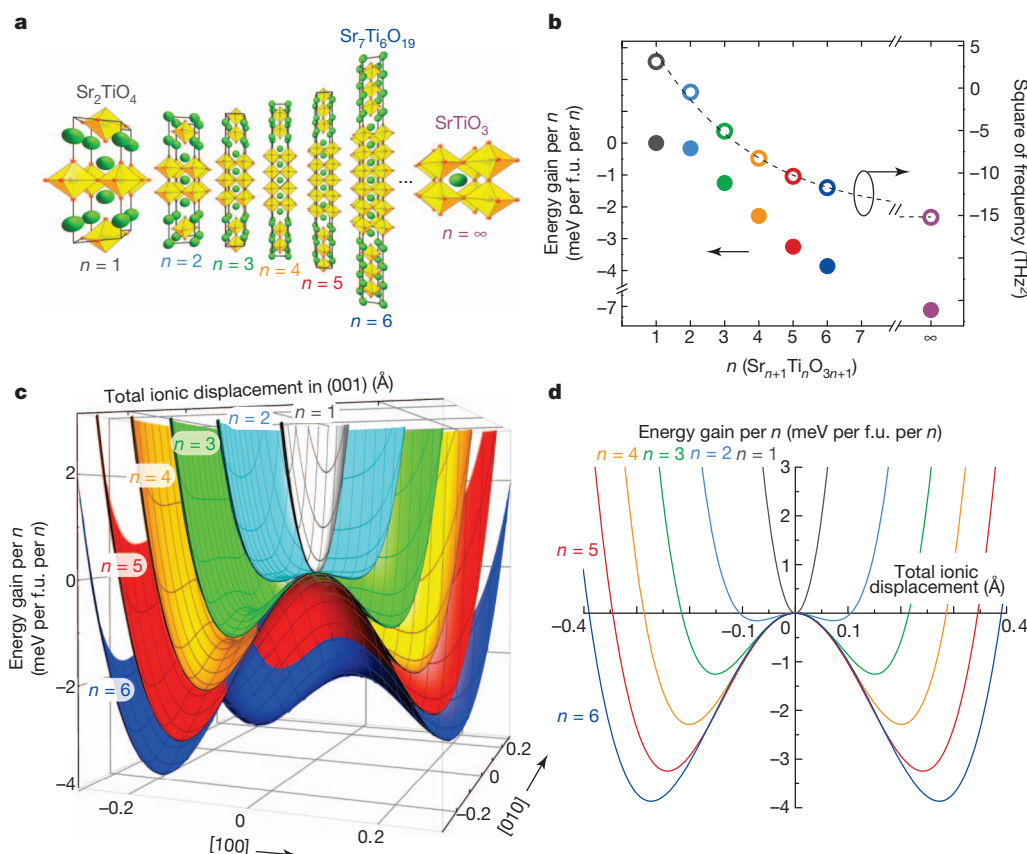


Figure 1 | First-principles calculations showing how the index n of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases strained commensurately to (110) DyScO_3 substrates can be used to control the local ferroelectric instability. **a**, Diagram of the crystal structure of a unit cell of the $n = 1$ – 6 and $n = \infty$ members of the $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases. **b**, Square of polar soft-phonon-mode frequency (right-hand axis) and the energy gain per n (left-hand axis) of the ferroelectric state with respect to the nonpolar state, calculated from first principles. Energy

be single-phase and commensurately strained to the substrates on which they were grown. The θ – 2θ XRD scan of each $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ film shows all expected peaks and seems to imply perfect layer periodicity along the out-of-plane direction in each sample. Bright-field scanning transmission electron microscope (STEM) images of the $n = 6/\text{DyScO}_3$ sample (Fig. 2b), however, show that there are not only periodic horizontal $(\text{SrO})_2$ planes but also aperiodic vertical $(\text{SrO})_2$ planes. A histogram analysis (Fig. 2c) of the layering disorder reveals that most of the layers along the growth direction are composed of the desired six perovskite layers. The remaining layers have spacings that are harmonics of $n = 6$ —that is, locally $n = 12, 18$ and 24 —that are well lattice-matched to the surrounding $n = 6$ matrix. These harmonic n values and vertical $(\text{SrO})_2$ layers probably form to accommodate local stoichiometry variations encountered during growth^{8,9}. Atomic models of the $n = 6$ phase illustrating its ability to accommodate local non-stoichiometry are provided in Fig. 2d–g.

The dielectric properties of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ samples were measured over a frequency range of 1 kHz to 125 GHz with a broadband, on-wafer technique²¹. The temperature dependence of the real part of the in-plane dielectric constant, $K_{11}(T)$, in the low-frequency regime (from 10 kHz to 1 MHz) is shown in Fig. 3a. The strong peak in $K_{11}(T)$, for $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ films with $n \geq 3$, is indicative of a phase transition from a paraelectric state above the transition temperature (T_C) to a state with local ferroelectric order below T_C , in agreement with theory (Fig. 1b–d).

The relationship between T_C and n determined from $K_{11}(T)$ measurements on films subjected to two different strain states by growth on DyScO_3 and GdScO_3 substrates is shown in Fig. 3b. T_C of the $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ ($n = 3$ – $6, \infty$) phases on both substrates systematically

gains are calculated by performing ionic relaxations under boundary conditions of fixed biaxial strain. F.u., formula unit. **c**, Energy (per n) with respect to the nonpolar state of $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ phases for polar distortions in the (001) plane. The distortion pattern is obtained from the ionic relaxations for larger values of n , and from the force constants matrix for smaller values of n (Methods). **d**, Cut of energy surfaces in **c** for polarization along the [110] axis.

increases with n , as expected from theory (second-harmonic generation (SHG) shows the same trend in T_C ; Extended Data Fig. 5). Also consistent is the higher T_C observed on GdScO_3 as a result of the larger strain. Figure 3c shows the dependence of the spontaneous polarization measured at 10 K on n for the films deposited on DyScO_3 .

Because the dielectric properties at room temperature are particularly important for applications, we examine the $n = 6/\text{DyScO}_3$ sample in detail at 300 K. Figure 4a shows the real and imaginary parts of the in-plane dielectric constant measured from 1 kHz to 125 GHz, demonstrating low loss and dispersionless response over almost the entire radiofrequency and microwave range. Even at ~ 1 THz, K_{11} remains unchanged (Methods), showing that phonons are the predominant contributor to K_{11} at 300 K over this broad frequency range. The inset to Fig. 4a shows the loss tangent of the same film on a linear frequency scale, indicating that only at the highest measurement frequencies does the loss become appreciable. Figure 4b shows the tunability at room temperature, indicating roughly 20% film tuning for a bias field of 50 kV cm^{-1} across the entire microwave range. On the basis of a model of the frequency dependence of the loss (Methods), we fit this loss tangent to a linear frequency dependence and calculate the film quality factor ($Q = 1/\tan\delta$), plotted in Fig. 4c. Also shown as solid symbols in Fig. 4c is the film Q , calculated by averaging the loss tangent data over a frequency window of width ~ 14.5 GHz. We then determine the film's FOM by multiplying the film Q by the relative tuning of 20%, obtaining the result shown in red in Fig. 4c. For comparison we also plot in Fig. 4c the best reliable report of the FOM of a $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ film at room temperature³⁰. Even though the FOM of the $\text{Ba}_x\text{Sr}_{1-x}\text{TiO}_3$ film is measured at a bias field sixfold higher, the $n = 6/\text{DyScO}_3$ sample

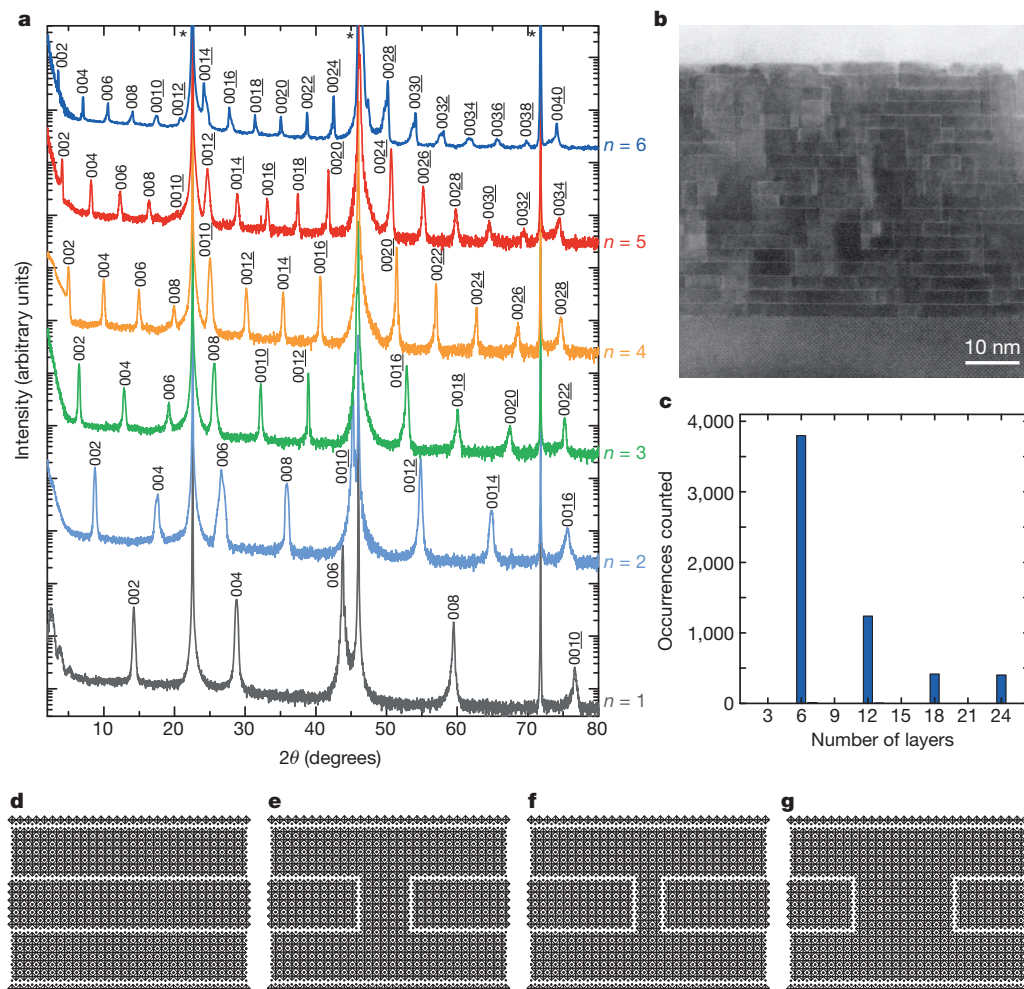


Figure 2 | Structural characterization by XRD and TEM.

a, θ - 2θ scans of epitaxial $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ ($n=1$ – 6) films grown on (110) DyScO_3 . Substrate peaks are labelled with an asterisk, and the plots are offset for clarity. **b**, Bright-field STEM image of the $n=6$ film grown on DyScO_3 . **c**, Histogram of the number of perovskite layers between $(\text{SrO})_2$ layers along the out-of-plane direction. **d**, Schematic two-dimensional atomic representation of the strontium atoms and TiO_6 octahedra of the ideal, stoichiometric $n=6$ phase. **e**, When vertical $(\text{SrO})_2$ layers are introduced, their overall density determines the SrO content of the film. This content is proportional to the length of the $(\text{SrO})_2$ crystallographic shear planes, which show up as white lines in this two-dimensional diagram. For the case drawn, the stoichiometry is the same as in **d**. **f**, **g**, Local non-stoichiometry accommodation is shown for the case of regions that are $\sim 7\%$ Sr-rich (**f**) and $\sim 12\%$ Sr-poor (**g**).

shows a significantly better FOM over the entire microwave frequency range, achieving a value of ~ 50 at 10 GHz. This $n=6$ sample has a higher FOM than any known electronically tunable dielectric at room temperature and comparable electric field.

The temperature dependence of the broadband frequency-dependent dielectric constant function of the $n=6/\text{DyScO}_3$ sample allows us to conclude that losses in this material are almost entirely due to polar nanoregions that have a finite distribution of sizes (Methods). These

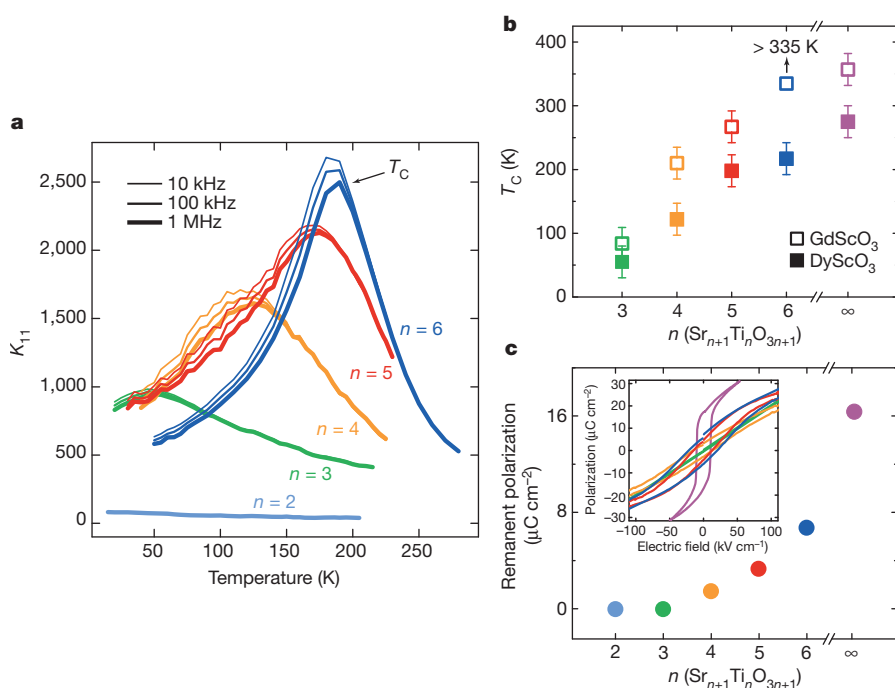


Figure 3 | Emergence of ferroelectricity in $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$ films grown on (110) DyScO_3 and (110) GdScO_3 . **a**, Temperature dependence of the real part of the in-plane dielectric constant (K_{11}) of $n=2$ – 6 films deposited on (110) DyScO_3 at 10 kHz, 100 kHz and 1 MHz. **b**, T_C as a function of n for the $n=3$ – 6 and $n=\infty$ films on (110) DyScO_3 and (110) GdScO_3 . T_C is taken to be the temperature at which K_{11} is greatest at a measurement frequency of 1 MHz. T_C of the $n=6/\text{DyScO}_3$ sample is indicated in **a**. The error bars correspond to the average variation of T_C among separately grown and measured samples having duplicate n values. The samples grown on DyScO_3 had duplicates, but not the samples on GdScO_3 . **c**, Remanent polarization at 10 K as a function of n for the $n=2$ – 6 and $n=\infty$ films on (110) DyScO_3 . The inset is a plot of polarization against electric field hysteresis loops measured at 10 K. The bright-field TEM image and the θ - 2θ rocking curve XRD scans from the same $n=6$ sample characterized in **a** are shown in Fig. 2b and Extended Data Fig. 3a, respectively.

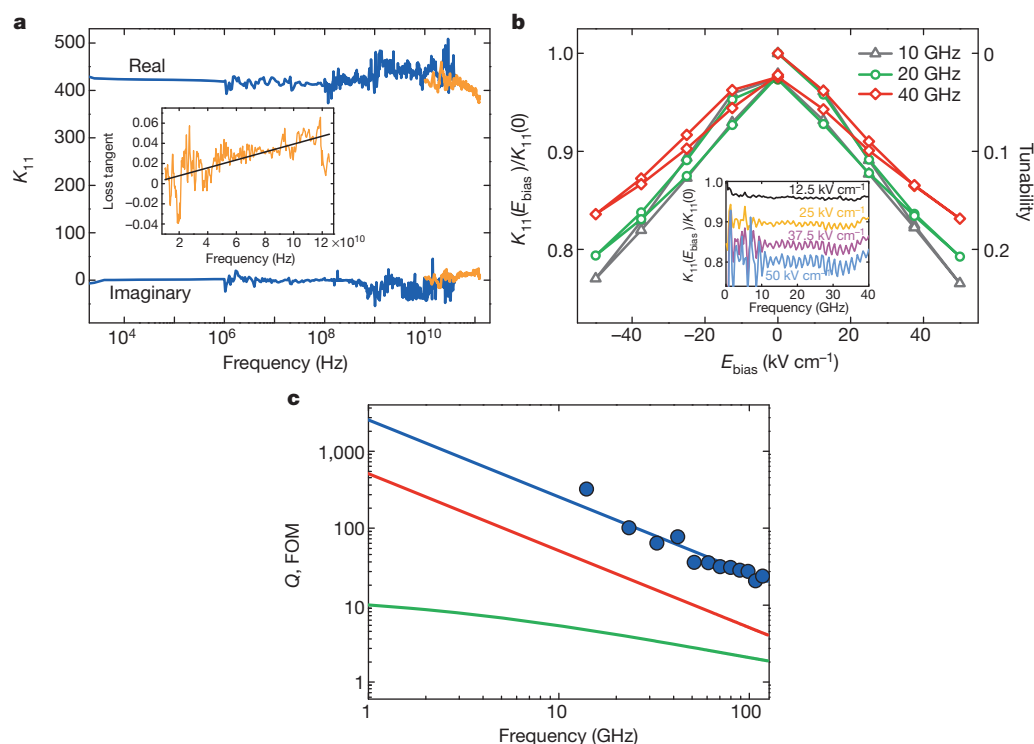


Figure 4 | In-plane dielectric constant (K_{11}) of $n = 6$ film on (110) DyScO₃, and its tunability at room temperature and high frequency. **a**, Real and imaginary parts of K_{11} as a function of frequency. Orange indicates the high-frequency results, measured on a linear frequency scale, from which the loss tangent is computed. The inset shows the film loss tangent on a linear frequency scale in the gigahertz frequency regime, along with the linear fit. **b**, The ratio of K_{11} under an applied bias field (E_{bias}) to that at zero bias field (left-hand axis) and tunability (right-hand axis) of the $n = 6$ sample at several different

frequencies in the microwave range. The inset shows the dielectric constant ratio as a function of frequency for several values of applied E_{bias} . **c**, Q_{11} (blue) and FOM (red) of the $n = 6$ sample at a bias field of 50 kV cm⁻¹, and the room-temperature FOM of a Ba_xSr_{1-x}TiO₃ film at 300 kV cm⁻¹ (green) from ref. 30. The FOM of the $n = 6$ sample assumes that the loss tangent depends linearly on frequency and that the tunability is independent of frequency and is 20% at a bias of 50 kV cm⁻¹. Solid points are Q values averaged over a frequency range of 14.5 GHz.

polar nanoregions nucleate far below room temperature; at those temperatures they enhance both K_{11} and loss, resulting in the observed dielectric relaxation typical for relaxor ferroelectrics. Such behaviour could arise from the horizontal (SrO)₂ planes decoupling the in-plane polarization between the perovskite slabs that they separate, leading to nanopolar slabs. Commensurate Sr_{*n*+1}Ti_{*n*}O_{3*n*+1} on DyScO₃ could thus be an embodiment of a new type of relaxor ferroelectric: one free of extrinsic disorder or the realization of a superparaelectric state. (There are also vertical (SrO)₂ planes. In contrast to the horizontal (SrO)₂ planes, however, the in-plane polarization of the Sr_{*n*+1}Ti_{*n*}O_{3*n*+1} film is perpendicular to the vertical (SrO)₂ planes. From electrostatic arguments analogous to those for BaTiO₃/SrTiO₃ superlattices³¹, the polarization should be continuous across the interface with the vertical (SrO)₂ planes.) The rapid decrease in the size of polar regions for $T > T_C$ is probably due to the lack of defects or the local nanostructure engineered into these materials, and is consistent with the exceptionally low dielectric loss and high FOM of the $n = 6$ film at room temperature in the microwave regime.

These results underscore the importance of both defect mitigation and our system that allows a ferroelectric instability to be tuned by means of atomic engineering without adding disorder. The high FOM already achieved extends the application of tunable dielectrics to significantly higher frequency. To allow this new material to be developed into practical devices, it is desirable to fabricate it on large-area substrates with low loss in the gigahertz frequency range. DyScO₃ substrates as large as 32 mm in diameter are currently grown by the Czochralski method³². These substrates could be scaled up, or the approach of making thick, relaxed DyScO₃ buffer layers (pseudo-substrates) on relevant low-loss substrates that are available in larger diameters could be pursued, as has been demonstrated for PrScO₃ (ref. 33). More broadly, however, a

multitude of other oxide systems, whose performance in thin-film form is limited by point defects, could also be greatly enhanced with appropriate atomic engineering. Exploiting host systems that form planar defects more readily than point defects is clearly advantageous.

METHODS SUMMARY

We performed first-principles density-functional calculations using Kohn–Sham density function theory as implemented in VASP and using density-functional perturbation theory as implemented in Quantum ESPRESSO (Methods and Extended Data Fig. 1). We grew Sr_{*n*+1}Ti_{*n*}O_{3*n*+1} ($n = 1$ –6) thin films by reactive molecular-beam epitaxy, from elemental strontium and titanium sources at a substrate temperature of 750–780 °C in an oxidant background pressure (O₂ + ~10% O₃) of 3×10^{-7} Torr (Methods). These films were characterized structurally by XRD (Extended Data Figs 2–4) and STEM. The paraelectric-to-ferroelectric transition was studied by SHG (Extended Data Fig. 5). The dielectric properties in the terahertz and infrared regime were measured by terahertz transmission and infrared reflectance (Methods and Extended Data Figs 6 and 7) and at microwave frequencies by on-wafer techniques with the use of interdigitated capacitors and coplanar waveguides (Methods and Extended Data Figs 8–10).

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 April; accepted 19 August 2013.

Published online 16 October 2013.

1. Tagantsev, A. K., Sherman, V. O., Astafiev, K. F., Venkatesh, J. & Setter, N. Ferroelectric materials for microwave tunable applications. *J. Electroceram.* **11**, 5–66 (2003).
2. Kirchoefer, S. W. *et al.* Microwave properties of Sr_{0.5}Ba_{0.5}TiO₃ thin-film interdigitated capacitors. *Microw. Opt. Technol. Lett.* **18**, 168–171 (1998).
3. Gevorgian, S. S. & Kollberg, E. L. Do we really need ferroelectrics in paraelectric phase only in electrically controlled microwave devices? *IEEE Trans. Microw. Theory Tech.* **49**, 2117–2124 (2001).

4. Nakamura, T. *et al.* On the perovskite-related materials of high dielectric permittivity with small temperature dependence and low dielectric loss. *Ferroelectrics* **196**, 205–209 (1997).
5. Wise, P. L. *et al.* Structure–microwave property relations in $(\text{Sr}_x\text{Ca}_{1-x})_{n+1}\text{Ti}_n\text{O}_{3n+1}$. *J. Eur. Ceram. Soc.* **21**, 1723–1726 (2001).
6. Andersson, S. & Wadsley, A. D. Crystallographic shear and diffusion paths in certain higher oxides of niobium, tungsten, molybdenum and titanium. *Nature* **211**, 581–583 (1966).
7. Anderson, J. S. *et al.* Point defects and extended defects in niobium oxides. *Nature* **243**, 81–83 (1973).
8. Tilley, R. J. D. Correlation between dielectric constant and defect structure of non-stoichiometric solids. *Nature* **269**, 229–231 (1977).
9. Tilley, R. J. D. An electron microscope study of perovskite-related oxides in the Sr–Ti–O system. *J. Solid State Chem.* **21**, 293–301 (1977).
10. Birol, T., Benedek, N. A. & Fennie, C. J. Interface control of emergent ferroic order in Ruddlesden–Popper $\text{Sr}_{n+1}\text{Ti}_n\text{O}_{3n+1}$. *Phys. Rev. Lett.* **107**, 257602 (2011).
11. Bao, P., Jackson, T. J., Wang, X. & Lancaster, M. J. Barium strontium titanate thin film varactors for room-temperature microwave device applications. *J. Phys. D Appl. Phys.* **41**, 063001 (2008).
12. Weiss, C. V. *et al.* Compositionally graded ferroelectric multilayers for frequency agile tunable devices. *J. Mater. Sci.* **44**, 5364–5374 (2009).
13. Pertsev, N. A., Tagantsev, A. K. & Setter, N. Phase transitions and strain-induced ferroelectricity in SrTiO_3 epitaxial thin films. *Phys. Rev. B* **61**, R825–R829 (2000).
14. Antons, A., Neaton, J. B., Rabe, K. M. & Vanderbilt, D. Tunability of the dielectric response of epitaxially strained SrTiO_3 from first principles. *Phys. Rev. B* **71**, 024102 (2005).
15. Li, Y. L. *et al.* Phase transitions and domain structures in strained pseudocubic (100) SrTiO_3 thin films. *Phys. Rev. B* **73**, 184112 (2006).
16. Haeni, J. H. *et al.* Room-temperature ferroelectricity in strained SrTiO_3 . *Nature* **430**, 758–761 (2004).
17. Cole, M. W., Nothwang, W. D., Hubbard, C., Ngo, E. & Ervin, M. Low dielectric loss and enhanced tunability of $\text{Ba}_{0.6}\text{Sr}_{0.4}\text{TiO}_3$ based thin films via material compositional design and optimized film processing methods. *J. Appl. Phys.* **93**, 9218–9225 (2003).
18. Cole, M. W., Joshi, P. C. & Ervin, M. H. La doped $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$ thin films for tunable device applications. *J. Appl. Phys.* **89**, 6336–6340 (2001).
19. Babbitt, R. W., Kosca, T. E. & Drach, W. C. Planar microwave electrooptic phase shifters. *Microwave J.* **35**, 63–79 (1992).
20. Gevorgian, S., Carlsson, E., Wikborg, E. & Kollberg, E. Tunable microwave devices based on bulk and thin film ferroelectrics. *Integr. Ferroelectr.* **22**, 765–777 (1998).
21. Booth, J. C., Takeuchi, I. & Chang, K. S. Microwave-frequency loss and dispersion in ferroelectric $\text{Ba}_{0.3}\text{Sr}_{0.7}\text{TiO}_3$ thin films. *Appl. Phys. Lett.* **87**, 082908 (2005).
22. Ruddlesden, S. N. & Popper, P. New compounds of the K_2NiF_4 type. *Acta Crystallogr.* **10**, 538–540 (1957).
23. Ruddlesden, S. N. & Popper, P. The compound $\text{Sr}_3\text{Ti}_2\text{O}_7$ and its structure. *Acta Crystallogr.* **11**, 54–55 (1958).
24. Lee, J. & Arias, T. A. Structural phase transitions in Ruddlesden–Popper phases of strontium titanate: *ab initio* and modulated Ginzburg–Landau approaches. *Phys. Rev. B* **82**, 180104 (2010).
25. Xi, X. X. *et al.* Oxide thin films for tunable microwave devices. *J. Electroceram.* **4**, 393–405 (2000).
26. Lee, C.-H. *et al.* Effect of reduced dimensionality on the optical band gap of SrTiO_3 . *Appl. Phys. Lett.* **102**, 122901 (2013).
27. Udayakumar, K. R. & Cormack, A. N. Structural aspects of phase-equilibria in the strontium–titanium–oxygen system. *J. Am. Ceram. Soc.* **71**, C469–C471 (1988).
28. Udayakumar, K. R. & Cormack, A. N. Non-stoichiometry in alkaline-earth excess alkaline-earth titanates. *J. Phys. Chem. Solids* **50**, 55–60 (1989).
29. Veličkov, B., Kahlenberg, V., Bertram, R. & Bernhagen, M. Crystal chemistry of GdScO_3 , DyScO_3 , SmScO_3 and NdScO_3 . *Z. Kristallogr.* **222**, 466–473 (2007).
30. Houzet, G., Burginies, L., Velu, G., Carru, J.-C. & Lippens, D. Dispersion and loss of ferroelectric $\text{Ba}_{0.5}\text{Sr}_{0.5}\text{TiO}_3$ thin films up to 110 GHz. *Appl. Phys. Lett.* **93**, 053507 (2008).
31. Neaton, J. B. & Rabe, K. M. Theory of polarization enhancement in epitaxial $\text{BaTiO}_3/\text{SrTiO}_3$ superlattices. *Appl. Phys. Lett.* **82**, 1586–1588 (2003).
32. Uecker, R. *et al.* Properties of rare-earth scandate single crystals (Re = Nd–Dy). *J. Cryst. Growth* **310**, 2649–2658 (2008).
33. Folkman, C. M., Das, R. R., Eom, C. B., Chen, Y. B. & Pan, X. Q. Single domain strain relaxed PrScO_3 template on miscut substrates. *Appl. Phys. Lett.* **89**, 221904 (2006).

Acknowledgements We acknowledge discussions with S. Trolier-McKinstry and C. A. Randall. Research was supported by Army Research Office (ARO) grants W911NF-09-1-0415 (for C.-H.L., Y.Z., J.A.M. and D.A.M.), W911NF-12-1-0437 (for Y.N., J.Z. and D.G.S.) and W911NF-10-1-0345 (for T.B., N.A.B. and C.J.F.); by the National Science Foundation (NSF) through Materials Research Science and Engineering Centers (MRSEC) grants DMR-0820404 (for R.H., E.V., X.X.X. and V.G.) and DMR-1120296 (for Y.K., J.D.B. and L.F.K.); by the Czech Science Foundation Project no. P204/12/1163 and the Czech Ministry of Education, Youth and Sports project LD12026 (for V.G., D.N. and S.K.); and by the Spanish Government and the European Union through grants EUI-ENIAC-2011-4349 and EUI-ENIAC 2010-04252 (for E.R.). C.-H.L. acknowledges stipend support from NSF grant DMR-0820404. J.A.M. acknowledges financial support from a National Defense Science & Engineering Graduate Fellowship. The dielectric and ferroelectric measurements in Fig. 3c were conducted at the Center for Nanophase Materials Sciences, which is sponsored at Oak Ridge National Laboratory by the Scientific User Facilities Division, Office of Basic Energy Sciences, US Department of Energy. This work was performed in part at the Cornell NanoScale Factory, a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation (grant ECCS-0335765). This work made use of the electron microscopy facility of the Cornell Center for Materials Research (CCMR) with support from the NSF MRSEC programme (DMR 1120296) and NSF IMR-0417392.

Author Contributions The thin films were synthesized by C.-H.L. on single-crystal substrates grown by M.B. and R.U. The first-principles calculations were performed by T.B., N.A.B. and C.J.F. The films were characterized by microwave measurements by N.D.O., E.R., I.T. and J.C.B.; by SHG by R.H., E.V. and V.G.; by infrared reflectance and terahertz transmission by V.G., D.N. and S.K.; by STEM by Y.Z., J.A.M., L.F.K. and D.A.M.; by XRD by C.-H.L., Y.N., J.-S.Z., Y.K. and J.D.B.; and by capacitance by M.D.B. and S.K. X.X.X. helped analyse the data. C.-H.L., T.B., J.C.B., C.J.F. and D.G.S. wrote the manuscript. The study was conceived and guided by C.J.F. and D.G.S. All authors discussed results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.G.S. (schlom@cornell.edu).

Robust twenty-first-century projections of El Niño and related precipitation variability

Scott Power¹, François Delage¹, Christine Chung¹, Greg Kociuba¹ & Kevin Keay¹

The El Niño–Southern Oscillation (ENSO) drives substantial variability in rainfall^{1–3}, severe weather^{4,5}, agricultural production^{3,6}, ecosystems⁷ and disease⁸ in many parts of the world. Given that further human-forced changes in the Earth’s climate system seem inevitable^{9,10}, the possibility exists that the character of ENSO and its impacts might change over the coming century. Although this issue has been investigated many times during the past 20 years, there is very little consensus on future changes in ENSO, apart from an expectation that ENSO will continue to be a dominant source of year-to-year variability^{9,11,12}. Here we show that there are in fact robust projected changes in the spatial patterns of year-to-year ENSO-driven variability in both surface temperature and precipitation. These changes are evident in the two most recent generations of climate models^{13,14}, using four different scenarios for CO₂ and other radiatively active gases^{14–17}. By the mid- to late twenty-first century, the projections include an intensification of both El-Niño-driven drying in the western Pacific Ocean and rainfall increases in the central and eastern equatorial Pacific. Experiments with an Atmospheric General Circulation Model reveal that robust projected changes in precipitation anomalies during El Niño years are primarily determined by a nonlinear response to surface global warming. Uncertain projected changes in the amplitude of ENSO-driven surface temperature variability have only a secondary role. Projected changes in key characteristics of ENSO are consequently much clearer than previously realized.

ENSO is a naturally occurring phenomenon centred in the equatorial Pacific arising from complex interactions between the atmosphere and ocean^{7,11,18}. Under climate change the possibility exists that the characteristics of ENSO will change during the coming century⁹. However, a recent review¹¹ concluded that ‘despite considerable progress in our understanding of the impact of climate change on many of the processes that contribute to El Niño variability, it is not yet possible to say whether ENSO activity will be enhanced or damped.’ This is consistent with the findings of an Intergovernmental Panel on Climate Change report⁹ and another recent review¹².

Nor is there a clear consensus on possible changes in ENSO teleconnections; that is, the impacts of ENSO outside the tropical Pacific. One study¹⁹ concluded that teleconnections that modulate the risk of drought will not change during the twenty-first century in climate models that simulate realistic twentieth-century ENSO variability. Another study²⁰ showed that wintertime ENSO teleconnections to the Northern Hemisphere shifted east under an idealized scenario of climate change in six of the eight models best able to simulate ENSO. This shift was attributed to an eastward shift in ENSO convection anomalies on the Equator by about 15° in longitude. However, the authors also concluded that an additional 14 models that less skilfully simulated ENSO did not show the same behaviour. In another study²¹ wintertime ENSO teleconnections to the North Pacific and North America were found to change in a consistent fashion in three different models, even though the models disagreed on the sign of projected change in ENSO amplitude. In recent times a new generation of climate models and scenarios has become available¹⁴, providing an opportunity to examine projected

changes in ENSO and ENSO-driven variability by using more models and more emission scenarios than ever before.

The multi-model average (MMA) difference between the twentieth-century (20C) and twenty-first-century (21C) standardized leading patterns of surface temperature (ST) variability associated with ENSO in these models (that is, $\Delta\text{EOF1} = \text{EOF1}(21\text{C}) - \text{EOF1}(20\text{C})$) is presented in Fig. 1. These patterns are derived from Empirical Orthogonal Function (EOF) analysis²². Results for four different twenty-first-century emission scenarios are presented (RCP8.5 in Fig. 1a, RCP4.5 in Fig. 1c, 1% CO₂ (where 1% CO₂ refers to a simple scenario in which greenhouse gases are increased by 1% per year (compounded)) in Fig. 1e, and SRES A2 in Fig. 1g). The corresponding twentieth-century EOFs are presented in Extended Data Fig. 1.

Even though many different models over two generations have been used, $\Delta\text{EOF1}(\text{TS})$ (Fig. 1a, c, e, g) shows a decrease near the Equator between 140° E and 165° E, and an increase between 170° W and 130° W under all four scenarios.

Changes in the standardized precipitation pattern ($\Delta\text{EOF1}(\text{precipitation})$; Fig. 1b, d, f, h) also show common changes in all four scenarios, with decreases in the west and increases in the central and eastern equatorial Pacific. Decreases are also evident around 7–10° N across most of the Pacific eastwards of 170° E, though the decreases are clearest and most common in the CMIP5 models under all three of the scenarios RCP8.5, RCP4.5 and 1% CO₂.

Composites of interannual (that is, ‘year-to-year’) ST anomalies in the twentieth and twenty-first centuries during El Niño years (as defined in Methods) are given for all four scenarios in Fig. 2. Agreement on ST change (Fig. 2a, c, e, g) is much less widespread near the Equator than it is for $\Delta\text{EOF1}(\text{ST})$ in Fig. 1. There are also marked differences in the projected changes between scenarios. For example, the 1% CO₂ scenario gives increases in the central equatorial Pacific with little agreement between the models on change in the western Pacific, whereas the RCP4.5 scenario shows a decrease in most models towards the west, and a MMA decrease in the central equatorial Pacific.

The greater extent of model agreement evident in $\Delta\text{EOF1}(\text{ST})$ (Fig. 1) compared with the extent of agreement in the El Niño ST composites (Fig. 2) can arise because the latter depends on changes in $\Delta\text{EOF1}(\text{ST})$ and on changes in the corresponding amplitude of variability associated with that pattern ($\Delta\eta$, say). Evidently there is less agreement between the models on $\Delta\eta$ than there is on $\Delta\text{EOF1}(\text{ST})$. This is confirmed in Extended Data Fig. 2, which shows that there is little consistency in amplitude changes between the scenarios. The proportion of models showing an increase in amplitude is 33% (RCP8.5), 48% (RCP4.5), 67% (1% CO₂) and 50% (A2).

Because tropical Pacific precipitation is a strong function of underlying ST^{7,18} and the ST changes are uncertain in many places (Fig. 2), it is very surprising to find widespread agreement among models and scenarios on precipitation changes over the equatorial Pacific (Fig. 2). This agreement includes an increase in precipitation near the Equator east of 170° E between roughly 5° N and 10° S, with decreases west of 150° E. There is also a decline near the Intertropical Convergence Zone² (ITCZ) around roughly 7–10° N in the CMIP5 models. This decline

¹Centre for Australian Weather and Climate Research, Bureau of Meteorology, Docklands, Melbourne, Victoria, 3008, Australia.

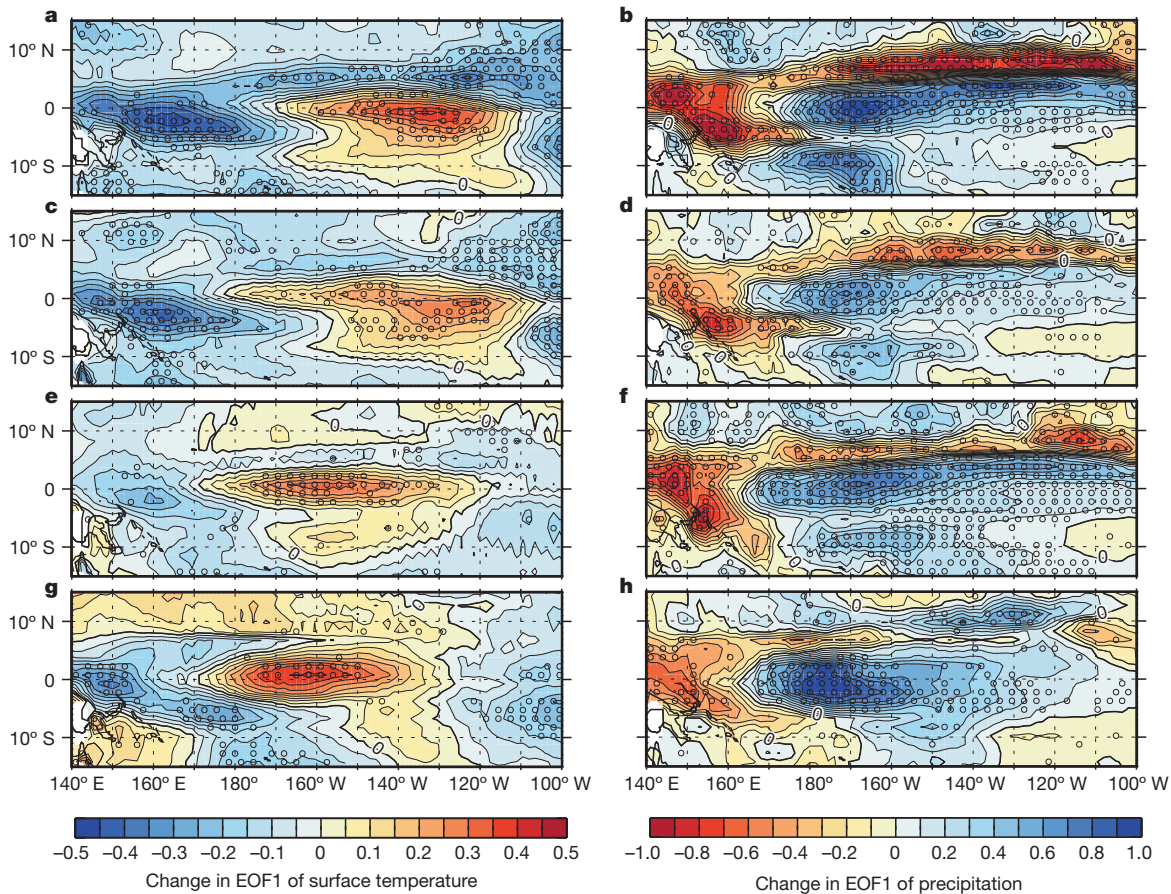


Figure 1 | Multi-model average (MMA) of the projected change in the structure of the standardized first EOF of interannual (high-pass-filtered, 'year-to-year') variability for the four twenty-first-century scenarios. a, c, e, g, Surface temperature (ST); b, d, f, h, precipitation. The pattern for each model was standardized by the spatial standard deviation of EOF1 over the

domain 0–360° E, 30° S to 30° N. The CMIP5 models were forced using RCP8.5 (a, b), RCP4.5 (c, d) and 1% CO₂ (e, f). The CMIP3 models were forced using SRES A2 (g, h). Stippling indicates that more than 70% of models agree on the sign of change. Red shades indicate an increase in EOF1 (ST) and a decrease in EOF1 (precipitation).

has been linked to an equatorward shift of the ITCZ in response to enhanced equatorial SST under global warming²³.

Somewhat the widespread uncertainty and marked differences between scenarios that are evident in ST variability do not lead to widespread uncertainty in projected changes in equatorial precipitation variability (Fig. 2). This is an important result because tropical precipitation has a vital role in ENSO dynamics^{7,18} and is a major driver of ENSO teleconnections^{20,21}. Atmospheric General Circulation Model (AGCM) experiments are conducted to investigate the cause of this apparent inconsistency. The experiments help to clarify the respective contributions of projected changes in the typical structure of El Niño sea surface temperature (SST) anomalies (SSTA), global warming and nonlinearity to the overall precipitation response. The AGCM is forced with several different spatially varying SSTA patterns (Extended Data Fig. 3), separately and combined.

Precipitation along the Equator in the AGCM is shown in Fig. 3a, b for three cases: experiment 20C gives the precipitation response to El Niño events of varying magnitude in the twentieth century; experiment 21C gives the same for the warmer twenty-first century under the assumption that the structure of SSTAs (relative to the new, warmer background climatology) associated with El Niño does not change; and experiment 21C + dSST gives the precipitation response in the twenty-first century allowing for projected changes in El Niño SSTAs in response to global warming.

Results for both RCP8.5 (Fig. 3a, c) and A2 (Fig. 3b, d) are consistent with the projected changes evident in the climate models (Fig. 2). For example, precipitation during El Niño (relative to the precipitation in

the corresponding case with no El Niño SSTA applied) tends to increase and shift towards the east, and to decrease towards the west. The results also show that the impact of global warming in the absence of structural changes to the El Niño SST anomaly (that is, in experiment 21C) depends strongly on the amplitude of the El Niño SSTA (α). For example, although global warming increases the precipitation response towards the east for all experiments with El Niño SSTAs applied, the magnitude of the precipitation change increases and the maximum response shifts further east as α increases.

The precipitation response to identical SSTAs in experiments 20C and 21C is different, indicating that the response is nonlinear because only the background-state (climatological) SSTs are different (see Methods for further details). The response is very similar in structure under the two scenarios. This consistency is due, in part, to the similarity of changes in the structure of mean-state SST across models and scenarios⁹ (Extended Data Figs 3b, c and 4).

The impact of the structural changes in the El Niño SST composite tends to decrease the impact of global warming on precipitation east of the dateline under the RCP8.5 scenario (Fig. 3c), whereas it tends to increase the impact under the A2 scenario (Fig. 3d). This contrast again reflects uncertainty in changes to ENSO-driven SST variability (Fig. 2 and Extended Data Fig. 3d, e) arising from uncertainty in amplitude changes. However, for every value of α the nonlinear response to unchanged El Niño SST tends to be either reinforced by or larger than the response to structural changes in El Niño SST (Fig. 3c, d). The dominance of the nonlinear response and its consistency across scenarios explains why there can be a consensus among the models and

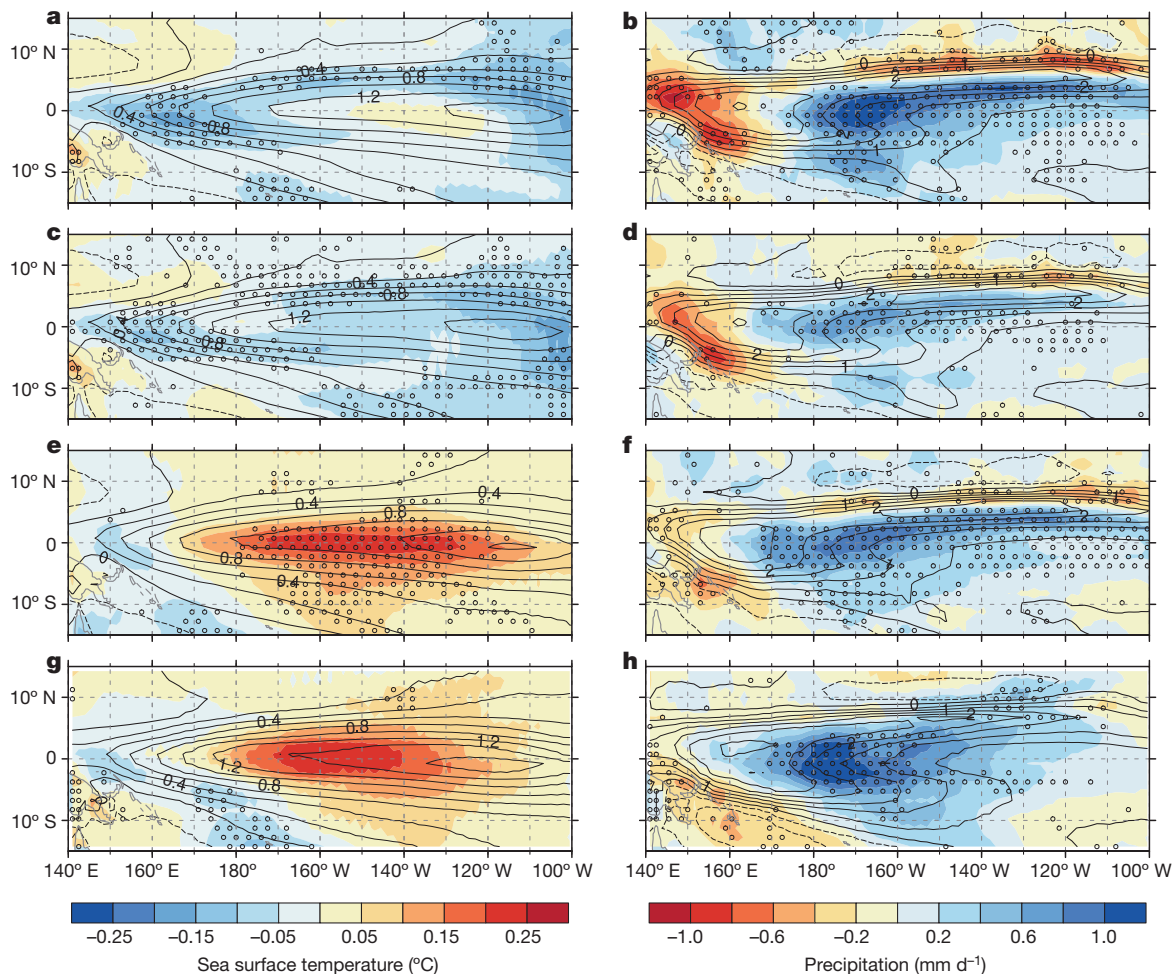


Figure 2 | MMA of the difference between twentieth-century and twenty-first-century filtered ST and precipitation anomalies in El Niño years. **a, c, e, g,** Surface temperature (ST); **b, d, f, h,** precipitation. **a, b,** RCP8.5. **c, d,** RCP4.5. **e, f,** 1% CO₂. **g, h,** The CMIP3 models were forced using SRES A2. The corresponding averages for El Niño–La Niña years are very similar

although larger in magnitude (not shown). The contour lines show the MMA of the twentieth-century anomalies during El Niño years. Stippling indicates agreement in more than 70% of models on the sign of change. Red shades indicate warming in ST, drying in precipitation.

scenarios on projected changes in ENSO-driven precipitation variability in the absence of a consensus on changes in ENSO-driven SST variability. The projected changes in El Niño precipitation anomalies (Figs 2b, d, f, h and 3) will tend to intensify and expand the El Niño-driven drying in the west Pacific (Fig. 3a, b), and reinforce El Niño-driven precipitation increases east of the dateline.

The nonlinear precipitation response in the AGCM depends partly on the distribution of ‘total SST’; that is, SST taking both mean-state changes and SST variability into account. To illustrate consistent behaviour in the coupled models, we examined the longitude of maximum equatorial SST. The maximum moves east over the period 1950–2099 in more than 75% of models. The shift exceeds about 4° east per century in 50% of models, and about 5.7° east per century in 25% of models. The maximum tends to move further east as the twenty-first century unfolds because of changes in the structure of EOF1 described above (Fig. 1) and because anthropogenic equatorial warming tends to be greater towards the east (Extended Data Fig. 4). As the maximum moves east during El Niño (because El Niño SSTA increases towards the east—see Extended Data Fig. 1), eastward shifting is enhanced if the ENSO amplitude increases, or is offset if the amplitude decreases. Any offsetting effect of decreased amplitudes in some models is evidently insufficient to reverse the eastward shifting from mean-state and EOF structural changes in more than 75% of models.

A previous study²⁴ showed that the frequency of ‘central Pacific El Niños’ increased relative to the frequency of ‘eastern Pacific El Niños’ in five of the six models best able to replicate the observed relative frequency. We find increases in the four CMIP5 models best able to simulate observations (see Methods section on central Pacific El Niños), lending support to previous conclusions²⁴. However, there is no consensus on such changes across all models. Because the changes highlighted in this investigation are evident in a much higher proportion of models, changes in the relative frequency of central Pacific El Niños do not account for the more robust changes that we identify.

Climate models are known to have systematic biases in their ability to simulate ENSO^{11,12,25}. One of the main problems is that the positive loading in model EOF1s tends to extend too far to the west (compare Extended Data Fig. 1a with Extended Data Fig. 1c, e). To examine the impact of this bias, each model’s twentieth-century EOF1 was shifted eastwards to give the best match to the observed pattern. The same shift was then applied to the corresponding twenty-first-century EOF1s. This simple correction leads to similar MMA projected changes to those obtained previously, with even larger changes and greater agreement between the models (Extended Data Fig. 5). This suggests that our major conclusions (summarized in Fig. 4) will apply to future generations of models with smaller biases.

The existence of robust projected changes in some other forms of ENSO-driven variability, while plausible, is an open question.

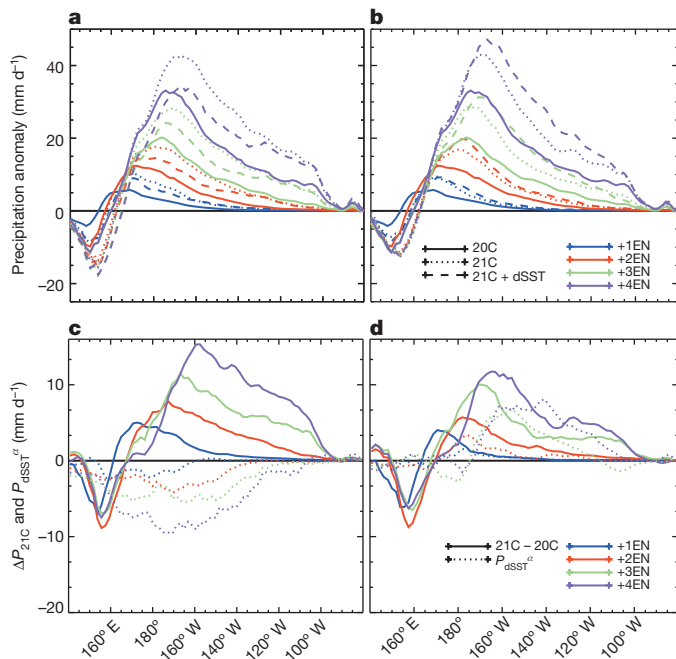


Figure 3 | Precipitation along the Equator over the Pacific in the AGCM. Experiments 1–5 as described in Extended Data Table 2. Values displayed in **a** and **b** are differences between precipitation with $\alpha \geq 1$ (that is, with α SSTA_EN applied) and the corresponding experiment with $\alpha = 0$. '1EN', for example, corresponds to the case $\alpha = 1$. '20C' values are differences relative to precipitation in the experiment with $\alpha = 0$ under twentieth-century conditions. '21C' values are differences relative to precipitation in the experiment with $\alpha = 0$ under twenty-first-century conditions (that is, with SSTA_GW applied). **a**, RCP8.5; **b**, A2. Values for the twentieth century (solid lines) and the twenty-first century with (dashed lines) and without (dotted lines) structural changes in the El Niño SST anomaly (Δ SSTA_EN) are shown in **a** and **b**. The key in **b** also applies to **a**. Values displayed in **c** and **d** show the impact of global warming (solid lines) and structural changes in the El Niño SST anomaly (dotted lines) on precipitation for $\alpha = 1, 2, 3$ and 4 ; **c**, RCP8.5; **d**, A2. ΔP_{21C} is the impact of global warming (that is, $21C - 20C$) on the El Niño precipitation anomaly, and P_{dSST}^{α} is the impact of structural change in the El Niño SST anomaly on the El Niño precipitation anomaly. The key in **d** also applies to **c**. See Methods for further details on the AGCM experiments conducted.

METHODS SUMMARY

Up to 21 models from the CMIP5 archive¹⁴ are used for the RCP8.5, RCP4.5 and 1% CO₂ experiments, and 16 models from CMIP3 (ref. 13) for the SRES A2 experiments. Details on the RCP and SRES scenarios are provided elsewhere^{14–16}. The full list of models used is given in Extended Data Table 1. All coupled climate models and the observations were re-gridded to a 1.5° latitude/1.5° longitude grid before analysis. A spectral filter was used to eliminate climate variability and changes with periods longer than 13 years. EOF analysis²² was used to extract the first ENSO pattern in the resulting interannual ST of every model. The resulting leading spatial pattern was then standardized by its spatial standard deviation (see Methods). The spatial standard deviations used to scale each EOF spatial pattern were also used to multiply the corresponding EOF time series to ensure that the product of the standardized pattern and the new time series remained equal to ST variability driven by the EOF (in kelvins). The EOF analysis was performed on June–December averages. The periods used were 1950–1999 and 2050–2099 (RCP8.5 and RCP4.5), 1939–1999 and 2038–2098 (A2) and years 1–50 and 51–100 (1% CO₂).

The AGCM was forced with several different spatially varying SSTA patterns (Extended Data Fig. 3): SSTA_GW, α SSTA_EN and α ΔSSTA_EN. SSTA_EN is the SST anomaly averaged over all El Niño events observed between 1978 and 2009, SSTA_GW is the MMA of the change in background SST projected for the twenty-first century, and ΔSSTA_EN is the MMA of the projected change in filtered SST during El Niño years. All experiments were conducted for $\alpha = 0, 1, 2, 3$ and 4 . The choice $\alpha = 0$ corresponds to climatological conditions, and $\alpha = 1, 2, 3$ and 4 correspond to weak, moderate, very strong and exceptionally strong El Niños, respectively.

Criteria used to define El Niño years are given in Methods.

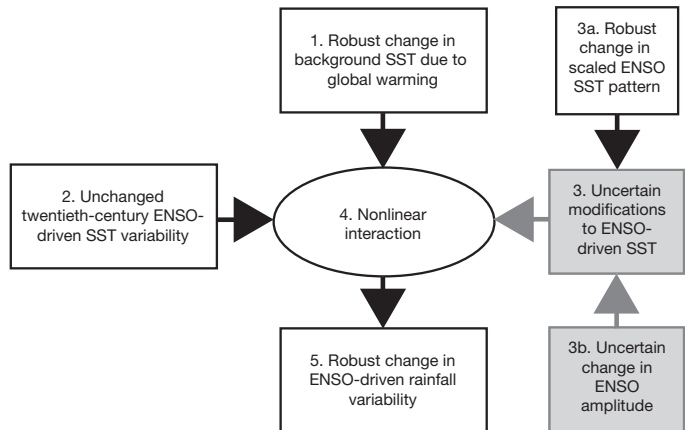


Figure 4 | Diagram illustrating main findings. It is often assumed that projected changes in ENSO amplitude are critically important in projections of all other ENSO impacts. We show instead that uncertainty in ENSO-driven SST variability does not necessarily exclude robust changes in other forms of ENSO-driven variability. This is apparent because robust projected changes in ENSO-driven rainfall variability in the equatorial Pacific are shown to occur despite uncertain changes in ENSO-driven SST variability. The rainfall response is caused by a nonlinear response to robust changes in background SST due to global warming (1), unchanged twentieth-century ENSO-driven SST variability (2), and uncertain changes in ENSO-driven SST variability (3). Components (1) and (2) dominate, resulting in more certainty in the rainfall projections than in the SST projections. The greater uncertainty in projected changes in ENSO-driven SST variability (3) results from uncertainty in projected change in the amplitude of SST variability (3a), and not from robust changes in the standardized pattern of ENSO-driven SST variability (3b). Grey shading is used to identify components making up the precipitation response that have uncertain projections.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 April; accepted 16 August 2013.

Published online 13 October 2013.

- Ropelewski, C. F. & Halpert, M. S. Precipitation patterns associated with the high index phase of the Southern Oscillation. *J. Clim.* **2**, 268–284 (2007).
- Australian Bureau of Meteorology and CSIRO. *Climate Change in the Pacific: Scientific Assessment and New Research* (eds Hennessy, K. Power, S. & Cambers, G.) (2011).
- Power, S., Casey, T., Folland, C., Colman, A. & Mehta, V. Interdecadal modulation of the impact of ENSO on Australia. *Clim. Dyn.* **15**, 234–319 (1999).
- Donnelly, J. P. & Woodruff, J. D. Intense hurricane activity over the past 5,000 years controlled by El Niño and the west African monsoon. *Nature* **447**, 465–468 (2007).
- Callaghan, J. & Power, S. B. Variability and decline in the number of severe tropical cyclones making land-fall over eastern Australia since the late nineteenth century. *Clim. Dyn.* **37**, 647–662 (2011).
- Hammer, G. L. *et al.* Advances in the application of climate prediction in agriculture. *Agric. Syst.* **70**, 515–553 (2001).
- McPhaden, M. J., Zebiak, S. E. & Glantz, M. H. ENSO as an integrating concept in earth science. *Science* **314**, 1740–1745 (2006).
- Sari Kovats, R. S. *et al.* El Niño and health. *Nature* **361**, 1481–1489 (2003).
- Meehl, G. A. *et al.* in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the 4th Assessment Report of the Intergovernmental Panel on Climate Change* (eds Solomon, S. *et al.*) 747–845 (Cambridge Univ. Press, 2007).
- Peters, G. P. *et al.* The challenge to keep global warming below 2°C. *Nature Clim. Change* **3**, 4–6 (2013).
- Collins, M. *et al.* The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geosci.* **3**, 391–397 (2010).
- Vecchi, G. A. & Wittenberg, A. T. El Niño and our future climate: where do we stand? *Wiley Interdisc. Rev. Clim. Change* **1**, 260–270 (2010).
- Meehl, G. *et al.* The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bull. Am. Meteorol. Soc.* **88**, 1383–1394 (2007).
- Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of the CMIP5 and the experimental design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
- Moss, R. H. *et al.* The next generation of scenarios for climate change research and assessment. *Nature* **463**, 747–756 (2010).
- van Vuuren, D. J. *et al.* The representative concentration pathways: an overview. *Clim. Change* **109**, 5–31 (2011).
- Nakicenovic, N. *et al.* *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2000).

18. Neelin, J. D. *et al.* ENSO theory. *J. Geophys. Res.* **103**, 14261–14290 (1998).
19. Coelho, C. A. S. & Goddard, L. El Niño-induced tropical droughts in climate change projections. *J. Clim.* **22**, 6456–6476 (2009).
20. Kug, J. S., An, S. I., Ham, Y. G. & Kang, I. S. Changes in El Niño and La Niña teleconnections over North Pacific–America in the global warming simulations. *Theor. Appl. Climatol.* **100**, 275–282 (2010).
21. Meehl, G. A. & Teng, H. Multi-model changes in El Niño teleconnections over North America in a future warmer climate. *Clim. Dyn.* **29**, 779–790 (2007).
22. Lorenz, E. N. *Empirical Orthogonal Functions and Statistical Weather Prediction* (Massachusetts Inst. Technology Sci. Rep. no. 1, 1956).
23. Xie, S.-P. *et al.* Global warming pattern formation: SST and rainfall. *J. Clim.* **23**, 966–986 (2010).
24. Yeh, S.-W. *et al.* El Niño in a changing climate. *Nature* **461**, 511–514 (2009).
25. Guilyardi, E. *et al.* Understanding El Niño in ocean–atmosphere general circulation models: progress and challenges. *Bull. Am. Meteorol. Soc.* **90**, 325–340 (2009).

Acknowledgements We thank L. Hanson, H. Ye and A. Moise for making the CMIP data available to us, and R. Colman, A. Handasyde and H. G. Power for comments on an earlier draft. This work is supported by the Australian Climate Change Science Program

(ACCSP) and the Pacific–Australia Climate Change Science and Adaptation planning Program (PACCSAP). PACCSAP is supported by AusAID, in collaboration with the Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led the development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Author Contributions S.P. drafted the paper, devised the hypotheses and was primarily responsible for designing the methods used. F.D., C.C., G.K. and K.K. contributed to method design, implemented the very complex code required, performed extensive data analyses, contributed to the writing of this paper, produced the plots and assisted in their interpretation. C.C. also conducted the AGCM experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.P. (s.power@bom.gov.au).

The origin of conodonts and of vertebrate mineralized skeletons

Duncan J. E. Murdock¹, Xi-Ping Dong^{2,3}, John E. Repetski⁴, Federica Marone⁵, Marco Stampanoni^{5,6} & Philip C. J. Donoghue¹

Conodonts are an extinct group of jawless vertebrates whose tooth-like elements are the earliest instance of a mineralized skeleton in the vertebrate lineage^{1,2}, inspiring the ‘inside-out’ hypothesis that teeth evolved independently of the vertebrate dermal skeleton and before the origin of jaws^{3–6}. However, these propositions have been based on evidence from derived euconodonts. Here we test hypotheses of a paraconodont ancestry of euconodonts^{7–11} using synchrotron radiation X-ray tomographic microscopy to characterize and compare the microstructure of morphologically similar euconodont and paraconodont elements. Paraconodonts exhibit a range of grades of structural differentiation, including tissues and a pattern of growth common to euconodont basal bodies. The different grades of structural differentiation exhibited by paraconodonts demonstrate the stepwise acquisition of euconodont characters, resolving debate over the relationship between these two groups. By implication, the putative homology of euconodont crown tissue and vertebrate enamel must be rejected as these tissues have evolved independently and convergently. Thus, the precise ontogenetic, structural and topological similarities between conodont elements and vertebrate odontodes appear to be a remarkable instance of convergence. The last common ancestor of conodonts and jawed vertebrates probably lacked mineralized skeletal tissues. The hypothesis that teeth evolved before jaws and

the inside-out hypothesis of dental evolution must be rejected; teeth seem to have evolved through the extension of odontogenic competence from the external dermis to internal epithelium soon after the origin of jaws.

The soft tissue anatomy of euconodonts substantiates their vertebrate affinity^{12,13}, but homology of euconodont and vertebrate skeletal tissues^{1,14,15} remains the subject of controversy^{16,17}. The mineralized skeleton of euconodonts consists of an oropharyngeal array of tooth-like elements that are composed of two mineralized structural elements, the crown and basal body which are comprised of tissues that resemble enamel and dentine⁸. Euconodont elements grew through centrifugal appositional growth, with laminae in the crown and basal body added in synchrony, in a manner comparable to enamel and dentine in the teeth of jawed vertebrates. However, knowledge of conodont skeletal tissues is based largely on extremely derived euconodonts and hypotheses of homology to canonical vertebrate skeletal tissues have taken no account of the evolutionary origin of the conodont skeleton. Based principally on similarities in morphology and patterns of growth, an evolutionary series was proposed originally among protoconodonts, paraconodonts and euconodonts^{7–11}. Protoconodonts have been recognized subsequently as stem-chaetognaths¹⁸ and excluded from euconodont ancestry, but the hypothesis that euconodonts are derived paraconodonts remains^{10,11}. Paraconodont elements are unipart, and have been considered

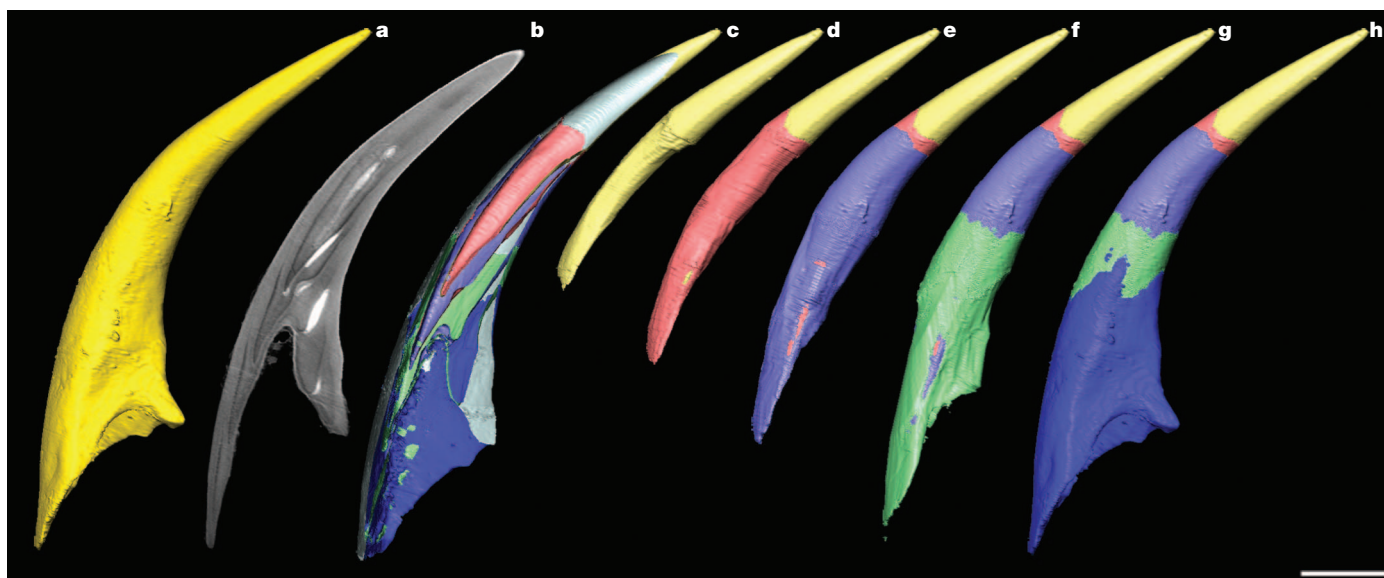


Figure 1 | Element growth and microstructure of the paraconodont *Furnishina*, Threadgill Creek section, Wilberns Formation, central Texas, 1,115 feet above base of Cambrian strata. a–c, The complete element has been subdivided into a number of discrete growth stages delimited by lines showing

cessation of growth (b, c). d–h, Initial growth stage, protoelement (d), is not enveloped by subsequent growth lamellae, rather lamellae are added to the proximal and lateral margins of the protoelement only (e–h). Scale bar, 50 μ m.

¹School of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, UK. ²School of Earth and Space Science, Peking University, Beijing 100871, China. ³State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China. ⁴US Geological Survey, MS 926A, National Center, Reston, Virginia 20192, USA. ⁵Swiss Light Source, Paul Scherrer Institut, Villigen 5232, Switzerland. ⁶Institute for Biomedical Engineering, University of Zürich and ETH Zürich, Rämistrasse, Zürich 8006, Switzerland.

homologous to the euconodont basal body alone because they grew through apposition of lamella layers to the proximal surface only. However, the histological comparisons of protoconodont and euconodont elements have been vague and aspects of paraconodont element structure and growth remain equivocal. For example, the homology of protoconodont elements and euconodont element basal bodies has been rejected on the basis that a basal body may not be primitive for euconodonts, and therefore could not be homologous to any paraconodont tissues^{19,20}. Key to the interpretation of paraconodont morphogenesis is the nature of the earliest stages of growth, or 'protoelement', which forms the distal-most part of the element. If characterized by complete centrifugal growth, this would result in a protoelement stage reminiscent of a euconodont crown plus basal body⁸. By contrast, addition of lamellae to the proximal surface only (that is, basal internal accretion) would result in a morphology reminiscent of the euconodont basal body alone. However, the evidential basis of this characterization has been criticized by some as an analytical artefact^{7,10,11}. We used synchrotron radiation X-ray tomographic microscopy (SRXTM) to characterize the element structure of paraconodonts and early euconodonts, non-invasively and at sub-micron resolution. We used the ensuing datasets to characterize the component tissues and uncover the pattern of development recorded in the sclerochronology of the growth arrest lines preserved in the mineralized tissues.

Based on the observed diversity of preserved structure we were able to divide paraconodont elements into three grades, each distinguished by the degree of tissue differentiation. Elements of *Furnishina* sp. exemplify the simplest grade of paraconodont elements. It consists of a single tissue type that exhibits punctuated incremental growth lines which define hollow conical laminae extending around the entire proximal margin and partly around the antapical margins (Fig. 1). Lamellae are oblique to the outer surface of the element and they do not extend over the distal tip, that is, the 'protoelement' is not enveloped by successive laminae (unlike the results in ref. 8). The basal cavity is not evident in the earliest laminae, rather developing in the latter stages; its depth is determined by the ontogenetic stage of development, for example, in elements of *Prooneotodus* sp., in earlier growth lamellae, resulting in a deeper basal cavity (Extended Data Fig. 1). The second grade of paraconodont element organization that we recognize is characterized by elements of *Problematoconites* sp., which is comprised of two tissues that have been identified previously as a distinct 'basal cone' and 'cone-filling'²¹. As in elements of *Furnishina*, the distal part of the element is formed of conical laminae, (basal cone of ref. 21). The proximal part of the element is formed from subsequent laminae extending across the entire proximal surface (cone-filling of ref. 21), forming a series of sub-parallel laminae—extensions of the laminae that comprise the rest of the element (Fig. 2). This is consistent with the model of a single secreting layer, (unlike the results in ref. 21). In our third grade of paraconodont element organization, exemplified by elements of *Rotundoconus tricarinatus* (Extended Data Fig. 2a), there are three principal tissue layers. The outermost layer consists of tapering rings that do not extend fully over the outer surface nor are they continuous over the proximal surface. These outer layers are bordered on the inside of the proximal surface by subparallel lamellae; it is unclear whether or not they converge at the apex. Finally, the basal cavity is filled with spheritic mineralization.

All euconodont elements exhibit a clear distinction between basal body tissue and crown tissue (for a guide to terminology see Extended Data Fig. 4). In the earliest euconodont elements, the basal body is indistinguishable from the most derived paraconodont elements. Following initial mineralization of the 'primordial element' subsequent laminae are added to the proximal margins. The basal body is differentiated into two tissue layers, distal hollow conical laminae and subparallel laminae across the proximal surface. These are formed from a single secreting layer (unlike the results in ref. 21). The crown tissue forms a cap over the entire surface of the basal body, thickening towards an enlarged cusp (Fig. 3). The relative size of the crown

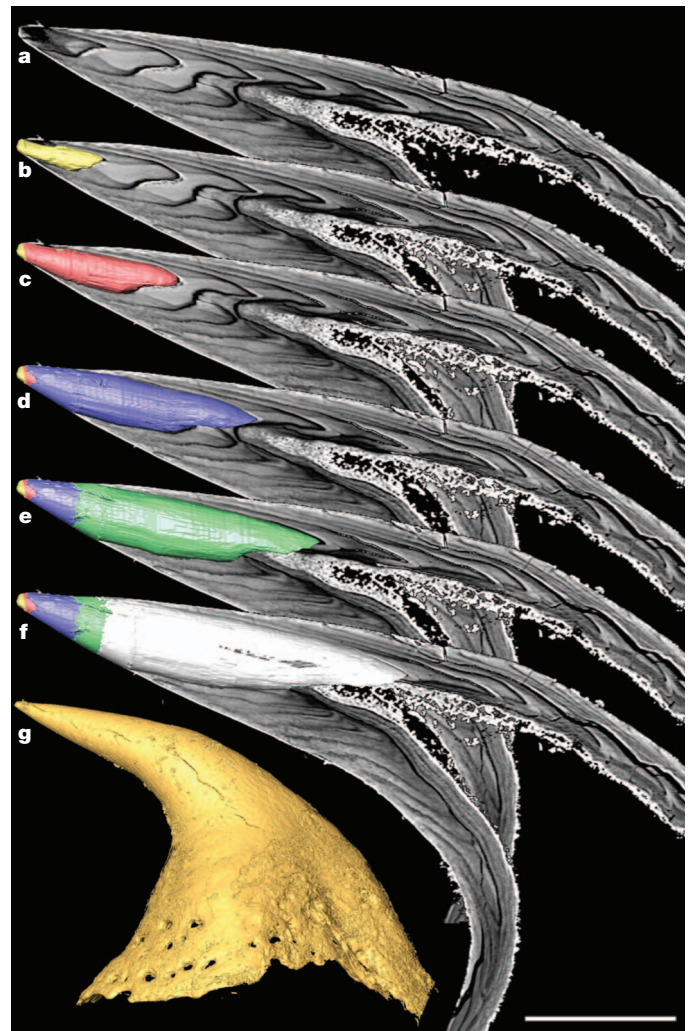


Figure 2 | Element growth and microstructure of the paraconodont *Problematoconites*, Windfall Formation, Tremadocian, Ordovician, Eureka County, Nevada, USA. a–g, Close-up of distal part of the cusp which has been subdivided into a number of discrete growth stages delimited by lines showing cessation of growth (b–f), with SRXTM rendering of complete element in the same orientation (g). Initial growth stage, protoelement, is not enveloped by subsequent growth lamellae, rather lamellae are added to the proximal and lateral margins of the protoelement only. Note the growth lamellae are continuous across the entire basal and margins of the element, not separated into basal cone and cone-filling (unlike the results in ref. 21). Scale bar represents 100 μ m (a–f); 266 μ m (g).

compared to the basal body is dictated simply by the degree to which the laminae of the crown extend beyond the distal tip of the basal body (compare elements of *Proconodontus serratus*; Extended Data Fig. 3, and *Proconodontus posterocostatus*; Fig. 3). White matter may be present in the crown (for example, in the posterior keel of the cusp of *P. serratus*; Extended Data Fig. 3). Other euconodont taxa retain the distinct three-layer structure of derived paraconodonts, for example, elements of *Granatodontus* sp. The entire element wall is thin and the basal cavity is deep (Extended Data Fig. 2b). A thin crown layer extends over the outer surface of the element, however, the basal body consists of two different tissues; a lamellar layer with sub-parallel lamellae surrounding a poorly defined porous tissue layer (Extended Data Fig. 2b).

Homology of the paraconodont element and the euconodont basal body was first proposed on the basis of simple observations of similarity in morphology and growth^{7,8,10,11}. However, these similarities have been insufficient to discriminate convergence from common descent. Our evidence reveals much greater complexity and differentiation in

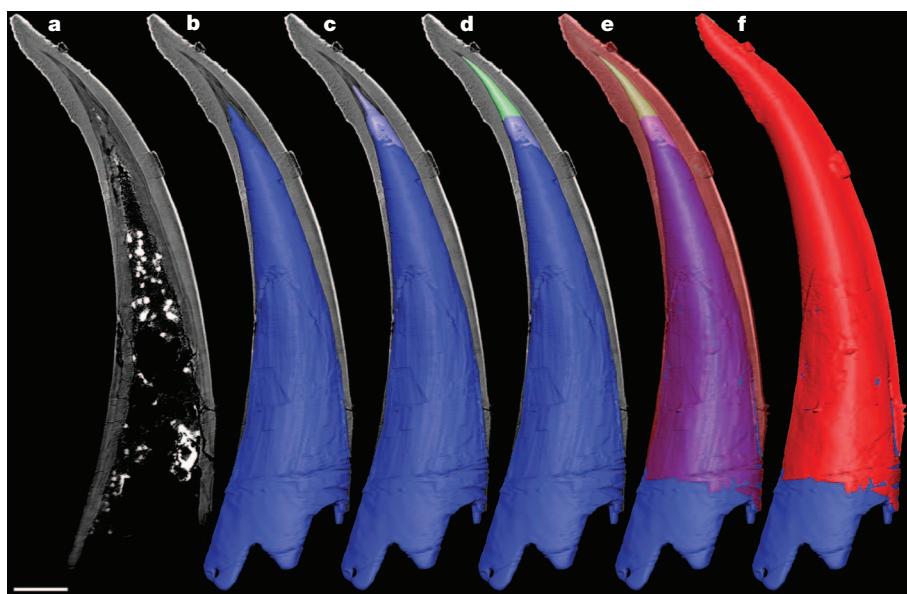


Figure 3 | Element growth of the euconodont *Proconodontus posterocostatus*, Gros Ventre Formation, Late Cambrian, Bighorn Mountains, Wyoming, USA. a, Longitudinal section showing delimitation of element into crown and basal body. b–f, SRXTM renderings of the initial two growth layers of basal body and the relationship between the crown (red) and basal body (blue, purple, green). The growth of the basal body continues as in elements of the paraconodont *Furnishina*, but with addition of crown tissue. Scale bar, 50 μ m.

the structure and growth of paraconodont elements than has been described previously, corroborating this hypothesis of homology. First, the protoelement of both the paraconodont element and euconodont element basal body is not overgrown at the distal tip. Rather, it is permanently exposed at the tip of protoconodont elements and remains in direct contact with the crown at the core of euconodont elements. Subsequent to the initial mineralization of the protoelement, the ontogeny of both paraconodont elements and the basal bodies of euconodont elements follow the same pattern of growth. The development of structural diversity exhibited by conodont elements is dictated simply by the relative timing of changes in the mode of secretion and, ultimately, through the differentiation of two principal structural elements, the basal body and crown, the latter characterizing the first euconodont elements. The basal cone and cone-filling structure described previously in euconodont basal bodies²¹ is manifest also in protoconodont elements, though we show that these are not separate structures and the growth lamellae are continuous between them. Crucially, the range of structures exhibited by the elements of different paraconodont species lie within nested sets of structural complexity, the most complex of which exhibit greater similarity to euconodont elements than other paraconodonts. Indeed, in terms of structure and arrangement of the component tissues, the basal bodies of elements of the early euconodont *Proconodontus* are effectively indistinguishable from the most complex paraconodont elements, such as those of *Problematoconites*. The same comparison can be made of the paraconodont *Rotundoconus* and the euconodont *Granatodontus*.

Direct comparison of ontogeny and tissue organization, coupled with a clear spectrum of complexity through early conodont elements, demonstrates that the similarities between paraconodont and euconodont elements go beyond analogy. Our results corroborate the hypothesis that the structural organization of the euconodont element was not only derived through the evolution of the enamel-like crown tissue from a paraconodont-grade ancestor, but also that characteristics of the euconodont basal body were assembled stepwise among different evolutionary grades of paraconodonts (Fig. 4). Evidently, the proposition of homology between euconodont crown tissue and vertebrate enamel^{11,15,22,23} fails a test of phylogenetic congruence²⁴ and must therefore be rejected. In this light, it is pertinent to question the proposed homology of euconodont basal tissue and vertebrate dentine since this is based largely on the topological and developmental relationship of euconodont basal tissue with crown tissue¹⁴. Among other early skeletonizing vertebrates, dentine is encountered only in the dermal skeleton, and it appears secondarily and convergently in the pharyngeal

and oral cavities of the jawless thelodonts²⁵ and early jawed vertebrates²⁶. Therefore there is no potential homologue of paraconodont elements in other total group gnathostomes. Thus, while it appears that conodonts afford the earliest manifestation of a mineralized skeleton in vertebrates, this skeleton evolved independently of other skeletonizing vertebrates. Although there is a remarkable similarity between euconodont elements and the odontodes of vertebrate scales and teeth, which extends from details of tissue microstructure through to the topological and developmental relationship among these tissues^{14,27}, it now appears to be a remarkable instance of evolutionary convergence. Euconodonts were influential in the hypothesis that teeth evolved before jaws and the ‘inside-out’ hypothesis in which dental evolution is independent of the tooth-like ‘odontode’ structures associated with external dermal scales^{3,4,6}. This view now lacks any evidential basis and must be rejected; teeth appear to have evolved through the

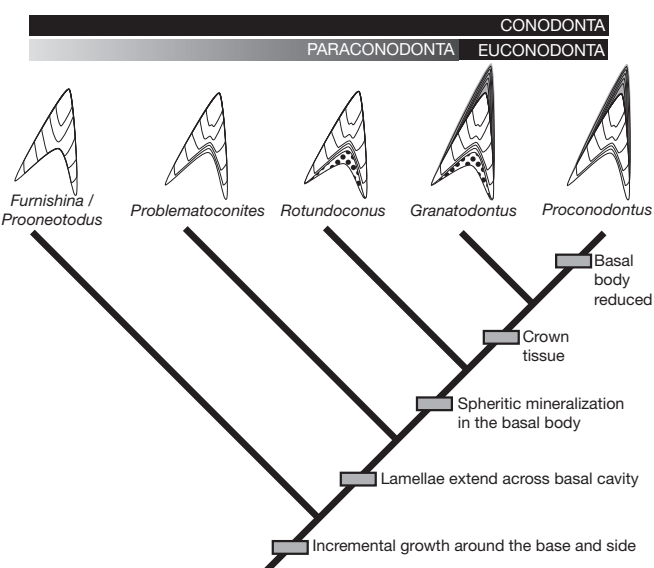


Figure 4 | Proposed phylogenetic hypothesis for the relationship between paraconodonts and euconodonts, and the evolution of conodont skeletal characters. Euconodonts are derived from a paraphyletic assemblage of paraconodonts that exhibit increasing basal body complexity, but are differentiated by the acquisition of the crown. Thus, the euconodont crown cannot be a homologue of vertebrate enamel.

extension of odontogenic competence from the external dermis to internal epithelium soon after the origin of jaws²⁶.

METHODS SUMMARY

We compared well-preserved, morphologically similar, paraconodont and euconodont elements from Middle Cambrian to Lower Ordovician age deposits; TC 1115, *Furnishina* sp. from Threadgill Creek section, Wilberns Formation, central Texas, 1,115 feet above base of Cambrian strata; USNM 593438, 593439 and 593440, *Prooneotodus* sp., *Problematoconites* sp., and *Proconodontus serratus* from the *Cambroistodus* subzone of the *Eoconodontus* zone of the Windfall Formation, Tremadocian, Ordovician, Eureka County, Nevada, USA; Lapworth Museum of Geology BU4421 *Proconodontus posterocostatus* from Gros Ventre Formation, Late Cambrian, Bighorn mountains, Wyoming, USA; GMPKU3068, *Rotundoconus tricarinaratus* from *Cordylodus intermedius* Zone, Furongian (Upper Cambrian), Panjiazui Formation, Wa'ergang section, Wa'ergangvillage, Taoyuan County, Hunan Province, China; USNM 521006, *Granatodontus* sp. from Steptoe South section, Whipple Cave Formation, uppermost Cambrian, northern Egan Range, White Pine County, Nevada, USA. Specimens were mounted on 3-mm brass stubs using clear nail varnish and volumetrically characterized using SRXTM²⁸. Measurements were taken using $\times 10$ and $\times 20$ objective lenses at 10–15 keV. For each data set, 1,501 projections over 180 degrees were acquired, resulting in volumetric data with voxel sizes of 0.74 and 0.36 μm , respectively. These experiments were performed on the TOMCAT beamline²⁹ at the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland. Figures were prepared using the VSG software Avizo (v6.4–7.1). Discrete growth stages or tissues, delimited by lines showing cessation of growth, were identified in the SRXTM slice data and individually labelled. These labels were then used to generate a three-dimensional surface representing the extent of an individual growth stage or tissue. Successive growth stages are distinguished by (arbitrary) colours.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 July; accepted 10 September 2013.

Published online 16 October 2013.

- Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. Presence of the earliest vertebrate hard tissues in conodonts. *Science* **256**, 1308–1311 (1992).
- Donoghue, P. C. J. & Sansom, I. J. Origin and early evolution of vertebrate skeletonization. *Microsc. Res. Tech.* **59**, 352–372 (2002).
- Smith, M. M. & Coates, M. I. Evolutionary origins of the vertebrate dentition: phylogenetic patterns and developmental evolution. *Eur. J. Oral Sci.* **106** (suppl. 1), 482–500 (1998).
- Smith, M. M. & Coates, M. I. in *Development, function and evolution of teeth* (eds Teaford M. F., Ferguson M. W. J., & Smith M. M.) 133–151 (Cambridge Univ. Press, 2000).
- Smith, M. M. & Coates, M. I. in *Major events of early vertebrate evolution* (ed. Ahlberg P. E.) 223–240 (Taylor & Francis, 2001).
- Fraser, G. J., Cerny, R., Soukup, V., Bronner-Fraser, M. & Streelman, J. T. The odontode explosion: the origin of tooth-like structures in vertebrates. *Bioessays* **32**, 808–817 (2010).
- Bengtson, S. Structure of some Middle Cambrian conodonts, and early evolution of conodont structure and function. *Lethaia* **9**, 185–206 (1976).
- Müller, K. J. & Nogami, Y. Über den Feinbau der Conodonten. *Memoirs of the Faculty of Science, Kyoto University, Series of Geology and Mineralogy* **38**, 1–87 (1971).
- Müller, K. J. & Nogami, Y. Growth and function of conodonts in *Proceedings of the 24th International Geological Congress* 20–27 (Montreal, 1972).
- Szaniawski, H. in *Palaeobiology of Conodonts* (ed. Aldridge R. J.) 35–47 (Ellis Horwood, 1987).
- Szaniawski, H. & Bengtson, S. Origin of euconodont elements. *J. Paleontol.* **67**, 640–654 (1993).
- Aldridge, R. J., Briggs, D. E. G., Smith, M. P., Clarkson, E. N. K. & Clark, N. D. L. The anatomy of conodonts. *Phil. Trans. R. Soc. Lond. B* **340**, 405–421 (1993).
- Pridmore, P. A., Barwick, R. E. & Nicoll, R. S. Soft anatomy and the affinities of conodonts. *Lethaia* **29**, 317–328 (1997).
- Donoghue, P. C. J. Growth and patterning in the conodont skeleton. *Phil. Trans. R. Soc. Lond. B* **353**, 633–666 (1998).
- Sansom, I. J., Smith, M. P. & Smith, M. M. Dentine in conodonts. *Nature* **368**, 591 (1994).
- Blieck, A. et al. Fossils, histology, and phylogeny: why conodonts are not vertebrates. *Episodes* **33**, 234–241 (2010).
- Turner, S. et al. False teeth: conodont-vertebrate phylogenetic relationships revisited. *Geodiversitas* **32**, 545–594 (2010).
- Szaniawski, H. Chaetognath grasping spines recognized among Cambrian protoconodonts. *J. Paleontol.* **56**, 806–810 (1982).
- Dzik, J. Remarks on the evolution of Ordovician conodonts. *Acta Palaeontol. Pol.* **21**, 395–453 (1976).
- Dzik, J. in *Problematic Fossil Taxa* (eds Hoffman A. & Nitecki M. H.) 240–254 (Oxford Univ. Press, 1986).
- Gross, W. Über die basis der Conodonten. *Palaeont. Zeits.* **31**, 78–91 (1957).
- Donoghue, P. C. J. Microstructural variation in conodont enamel is a functional adaptation. *Proc. R. Soc. Lond. B* **268**, 1691–1698 (2001).
- Donoghue, P. C. J., Purnell, M. A. & Aldridge, R. J. Conodont anatomy, chordate phylogeny and vertebrate classification. *Lethaia* **31**, 211–219 (1998).
- Patterson, C. in *Problems of Phylogenetic Reconstruction. Systematics Association Special Volume Vol. 29* (eds Joysey K. A. & Friday A. E.) 21–74 (Academic Press, 1982).
- Rücklin, M., Giles, S., Janvier, P. & Donoghue, P. C. J. Teeth before jaws? Comparative analysis of the structure and development of the external and internal scales in the extinct jawless vertebrate *Loganellia scotica*. *Evol. Dev.* **13**, 523–532 (2011).
- Rücklin, M. et al. Development of teeth and jaws in the earliest jawed vertebrates. *Nature* **491**, 748–751 (2012).
- Donoghue, P. C. J. & Aldridge, R. J. in *Major events in early vertebrate evolution: palaeontology, phylogeny, genetics and development* (ed. Ahlberg P. E.) 85–105 (Taylor & Francis, 2001).
- Donoghue, P. C. J. et al. Synchrotron X-ray tomographic microscopy of fossil embryos. *Nature* **442**, 680–683 (2006).
- Stampanoni, M. et al. Trends in synchrotron-based tomographic imaging: the SLS experience. *Proc. SPIE* **6318**, 63180M (2006).

Acknowledgements The SRXTM experiments were performed on the TOMCAT beamline at the Swiss Light Source, Paul Scherrer Institut (Villigen, Switzerland), funded through a project awarded to P.C.J.D. and S. Bengtson (Stockholm). NERC grant NE/G016623/1 to P.C.J.D., a studentship to DJEM funded by NERC and the Paul Scherrer Institut, and NSFC Project 41372015 to X.-P.D. Thanks to R. Stamm (USGS) for reviewing a draft of this manuscript; and thanks to J. E. Cunningham, D. O. Jones and M. Rücklin for assistance at the beamline. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author Contributions DJEM and P.C.J.D. conceived and designed the research; DJEM, F.M. and M.S. collected the SRXTM data; J.E.R. and X.-P.D. provided material and taxonomic information; DJEM analysed the data, prepared the figures and wrote the paper with substantive edits from P.C.J.D. and minor edits from the remaining authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.C.J.D. (phil.donoghue@bristol.ac.uk).

The maize *Ga* gene *COMPACT PLANT2* functions in *CLAVATA* signalling to control shoot meristem size

Peter Bommert¹, Byoung Il Je¹, Alexander Goldshmidt^{1†} & David Jackson¹

Shoot growth depends on meristems, pools of stem cells that are maintained by a negative feedback loop between the *CLAVATA* pathway and the *WUSCHEL* homeobox gene¹. *CLAVATA* signalling involves a secreted peptide, *CLAVATA3* (*CLV3*)², and its perception by cell surface leucine-rich repeat (LRR) receptors, including the *CLV1* receptor kinase³ and a LRR receptor-like protein, *CLV2* (ref. 4). However, the signalling mechanisms downstream of these receptors are poorly understood, especially for LRR receptor-like proteins, which lack a signalling domain⁵. Here we show that maize *COMPACT PLANT2* (*CT2*) encodes the predicted α -subunit (*Ga*) of a heterotrimeric GTP binding protein. Maize *ct2* phenotypes resemble *Arabidopsis thaliana clavata* mutants, and genetic, biochemical and functional assays indicate that *CT2/Ga* transmits a stem-cell-restrictive signal from a *CLAVATA* LRR receptor, suggesting a new function for *Ga* signalling in plants. Heterotrimeric GTP-binding proteins are membrane-associated molecular switches that are commonly activated by ligand binding to an associated seven-pass transmembrane (7TM) G-protein-coupled receptor (GPCR)⁶. Recent studies have questioned the idea that plant heterotrimeric G proteins interact with canonical GPCRs⁷, and our findings suggest that single pass transmembrane receptors act as GPCRs in plants, challenging the dogma that GPCRs are exclusively 7TM proteins.

The *ct2* reference allele (*ct2-Ref*) was obtained from the Maize Genetics Stock Center and introgressed into various inbred lines. It showed strong expressivity in B73, which we used for phenotypic characterization. *ct2* mutants displayed a range of phenotypes, including a shorter stature (Fig. 1a and Extended Data Fig. 1a, b), and shorter and wider leaves (Extended Data Table 1). *ct2* shoot apical meristems

(SAMs) were also wider, with an average diameter of 134 μm (± 6.8) compared to 109 μm (± 5.8 , $n = 15$; P value = 0.001; Student's t test) for normal siblings (Extended Data Fig. 1c, d). Despite the larger SAM, its identity and organization appeared normal in *ct2*, as shown by *in situ* hybridization with *KNOTTED1* (Extended Data Fig. 1e–h)⁸.

ct2 mutants also had striking inflorescence defects, including strongly fasciated ears (Fig. 1b, c) and thicker tassels (Fig. 1g, h), with a higher density of flower bearing structures known as spikelets (Extended Data Fig. 1k, l), resembling maize *CLAVATA* receptor mutants^{9,10}. To analyse inflorescence development, we used scanning electron microscopy (SEM). In wild type, the inflorescence meristem initiates spikelet pair meristems (SPMs) in a regular phyllotaxy, and SPMs branch to generate a pair of spikelet meristems in adjacent vertical rows, corresponding to rows of seeds in the cob (Fig. 1d and Extended Data Fig. 1i). When *ct2* ears were approximately 2 mm in length, the inflorescence meristem was enlarged (Fig. 1e and Extended Data Fig. 1j), leading to extra rows of SPMs, and meristem enlargement became more severe during development (Fig. 1f). *ct2* tassel meristems were also abnormally enlarged (Fig. 1i, j).

ct2 mapped to the short arm of Chromosome 1 (<http://www.maizegdb.org>), and using $\sim 1,000$ *ct2-Ref* F₂ mutants, we fine mapped it to a 1.2 megabase pair (Mbp) region containing approximately 30 genes. Among these was one encoding the α -subunit of a heterotrimeric GTP protein (Fig. 2a). Based on similar dwarf phenotypes in rice *Ga* mutants¹¹, we sequenced the locus from *ct2-Ref*, and found a 126 bp insertion within exon 14 (Fig. 2b). Three additional *ct2* alleles were isolated using a targeted ethylmethane sulphonate (EMS) screen, and each contained transition mutations in conserved splice sites, causing aberrant splicing

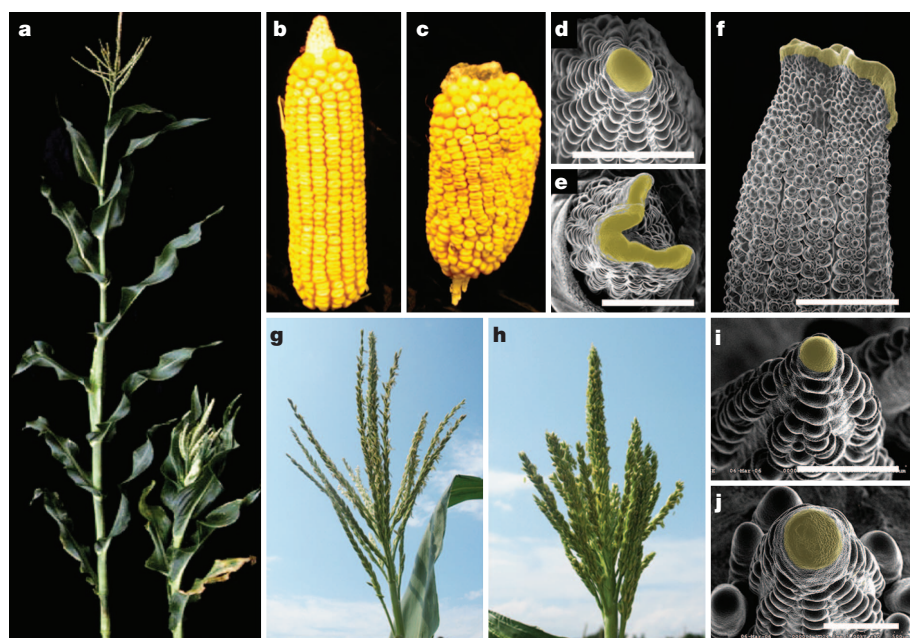


Figure 1 | *ct2* mutant phenotypes. **a**, *ct2* mutants (right) are semi-dwarfed, compared to wild-type sib. **b**, **c**, Wild-type (**b**) and *ct2* (**c**) ear showing fasciation. **d**, Top-down view of a wild-type ear primordium, the inflorescence meristem is shaded in yellow. **e**, *ct2* ear primordium showing enlargement and fasciation. **f**, Older *ct2* ear is more fasciated. **g**, Wild-type tassel. **h**, *ct2* tassel has as a thicker appearance due to increased spikelet density. **i**, Top-down view of a wild-type tassel inflorescence meristem (shaded). **j**, *ct2* tassel shows enlarged inflorescence meristem (shaded). Scale bars represent 1 mm (**d**, **e**), 2 mm (**f**) and 500 μm (**i**, **j**).

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. [†]Present address: Monsanto Company, 700 Chesterfield Parkway, Chesterfield, Missouri 63017, USA.

(Fig. 2b and Extended Data Fig. 2a). Each mutation was predicted to introduce premature stop codons, suggesting they are null alleles. Our characterization of four independent alleles indicates that *ct2* encodes the $G\alpha$ subunit of a predicted heterotrimeric G protein.

We next expressed a CT2 fusion with yellow fluorescent protein (YFP) driven by the CT2 endogenous promoter. This fusion complemented *ct2* mutants (Extended Data Fig. 2b–f), and CT2–YFP was observed in a thin line at the cell periphery that co-localized with an FM4-64 plasma membrane counterstain after plasmolysis (Fig. 2c, d). CT2–YFP also showed co-localization with FM4-64 in the SAM (Fig. 2e), and was expressed throughout the SAM and developing leaf primordia (Fig. 2f, g), and in the inflorescence meristem, where it was enriched in the outer layers (Fig. 2h and Extended Data Fig. 3a). Expression persisted throughout spikelet and floral development (Extended Data Fig. 3b–e). We also detected CT2–YFP expression in roots, again along the cell periphery, consistent with its predicted plasma membrane localization (Extended Data Fig. 3f). In summary, CT2–YFP appeared to localize to plasma membranes in meristems and in developing organs.

As the *ct2* inflorescence phenotypes were reminiscent of other maize fasciated ear mutants^{9,10}, we analysed genetic interactions between *ct2* and *thick tassel dwarf1* (*td1*) or *fasciated ear2* (*fea2*), which encode maize orthologues of *CLV1* and *CLV2*, respectively. To obtain a quantitative measure of phenotypic strength, we counted spikelet density in double mutant segregating families, as in other studies^{9,10}. Spikelet density of *ct2-Ref*; *td1-Ref* double mutants was significantly higher than either single mutant (P value = 0.0001; Student's t test), indicating an additive genetic effect and suggesting *ct2* and *td1* act in different pathways. In contrast, spikelet density in *ct2-Ref*; *fea2-0* double mutants was not significantly different from that of the *ct2* single mutants (P value = 0.42; Student's t test), even though each single mutant was significantly higher than normal (P value = 0.001; Student's t test) (Extended Data Fig. 4). This genetic interaction suggests they act in a common pathway. To substantiate these findings, we measured SAM diameter in *ct2-Ref*; *fea2-0* double mutant segregating families. In the double mutants, SAM diameter was not significantly different from that of *fea2-0* single mutants (P value = 0.41; Student's t test), even though each single mutant was significantly higher than normal (P value = 0.001; Student's t test) (Fig. 3a). This genetic interaction indicates that *fea2* is epistatic to *ct2* with respect to SAM diameter, and suggests they act in a common pathway. However, *ct2* mutant meristems were significantly smaller than the double mutants (Fig. 3a), suggesting that FEA2 signals through other pathways in addition to CT2/ $G\alpha$ to control SAM size.

To investigate the molecular basis for the epistatic interaction, we made a peptide antiserum against FEA2, and used a GFP antiserum to detect CT2–YFP. The anti-FEA2 antiserum detected a protein with an apparent weight of ~75 kDa, slightly increased compared to its predicted 61 kDa, in extracts from wild type (B73) but not *fea2* mutant ears, indicating that it was specific (Fig. 3b). Treatment with PNGaseF showed that FEA2, like many other LRR receptors, is glycosylated¹², and aqueous two-phase partitioning showed that it is predominantly present in the plasma membrane (Extended Data Fig. 5b, c). We also detected a band of the predicted size in total, soluble and membrane enriched extracts from CT2–YFP plants, but not from non-transgenic (B73) plants, and CT2–YFP was enriched in membrane fractions (Fig. 3c and Extended Data Fig. 5d). FEA2 and CT2 were found in overlapping higher molecular weight native complexes (Extended Data Fig. 6), and we used extracts from the CT2–YFP maize lines for immunoprecipitation experiments. Membrane-enriched extracts were immunoprecipitated using anti-GFP antiserum and following stringent washing the immunoprecipitate was probed by western blotting. We detected FEA2 in the input, and also in the immunoprecipitated fraction, suggesting that FEA2 interacts with CT2 (Fig. 3d and Extended Data Fig. 5e). We used a different membrane-localized fusion protein, PIN1–YFP¹³, as a control, and FEA2 was not detected in the immunoprecipitated fraction, nor was it detected in controls using wild type B73 extracts

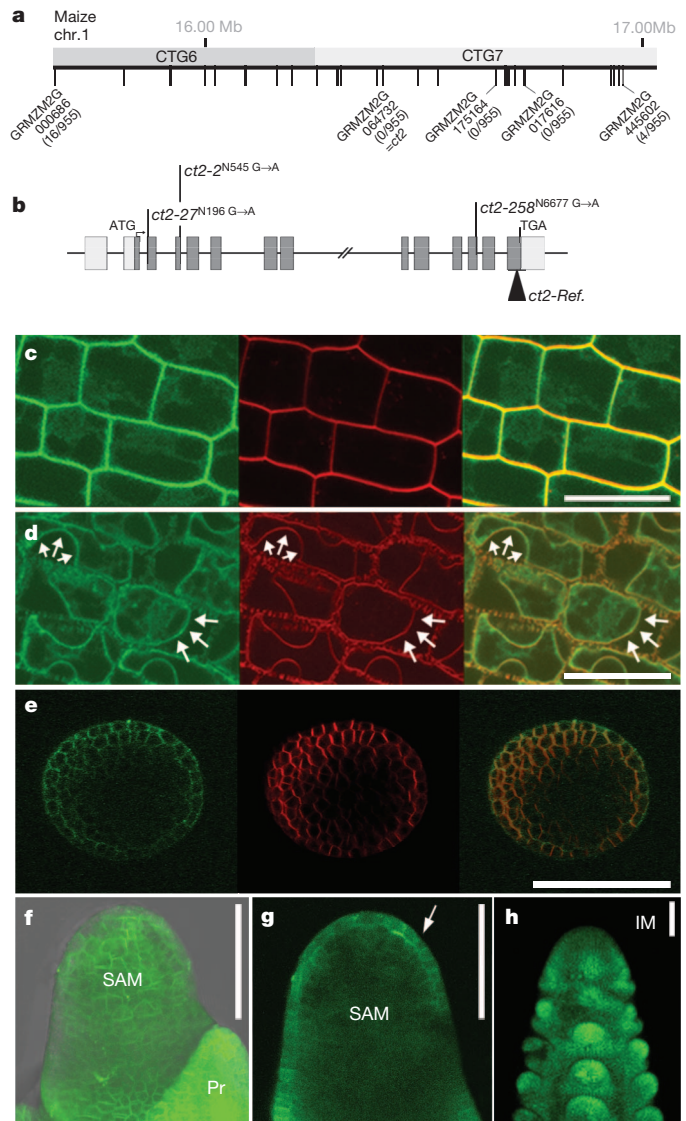


Figure 2 | Cloning and expression of CT2. **a**, Positional cloning of *ct2*. Numbers of recombinants/ F_2 individuals indicated. Vertical lines represent gene models. **b**, *ct2* gene, exons shown as grey boxes and different alleles are marked. **c**, Leaf cells expressing CT2–YFP (green), counterstained with FM4-64 (red), visible as a thin line around the cell. Overlay (right). **d**, Following plasmolysis, CT2–YFP (arrows) remains co-localized with FM4-64. **e**, Two-photon images of SAM expressing CT2–YFP, counterstained with FM4-64. **f**, CT2–YFP expression throughout the SAM and leaf primordium (Pr). **g**, Confocal section shows L1 layer enrichment. **h**, CT2–YFP in immature ears. Three-dimensional reconstructions are shown in **f** and **h**. Scale bars, 100 μ m.

(Fig. 3d and Extended Data Fig. 5e). These data indicate that CT2 interacts specifically with FEA2 *in vivo*, and together with their epistatic genetic interaction supports the hypothesis that CT2 signals in a FEA2 receptor pathway.

Our data suggest that CT2/ $G\alpha$ functions in a CLAVATA signalling pathway. To further test this hypothesis, we asked if *ct2* mutants show altered sensitivity to the CLV3 ligand. CLV3 function can be assessed by adding exogenous peptide, which inhibits meristem growth¹⁴. Maize seedling shoot meristems are covered by leaf primordia, so we used embryos, where the SAM is exposed. Wild-type embryos grew normally in culture in the presence of a control, scrambled CLV3 peptide (sCLV3), but SAM growth was strongly inhibited in the presence of CLV3 (Fig. 3e). *ct2* mutant embryos also grew normally in the presence of sCLV3, but showed a significantly reduced sensitivity to CLV3 (Fig. 3e), supporting the idea that CT2 is involved in transmitting a

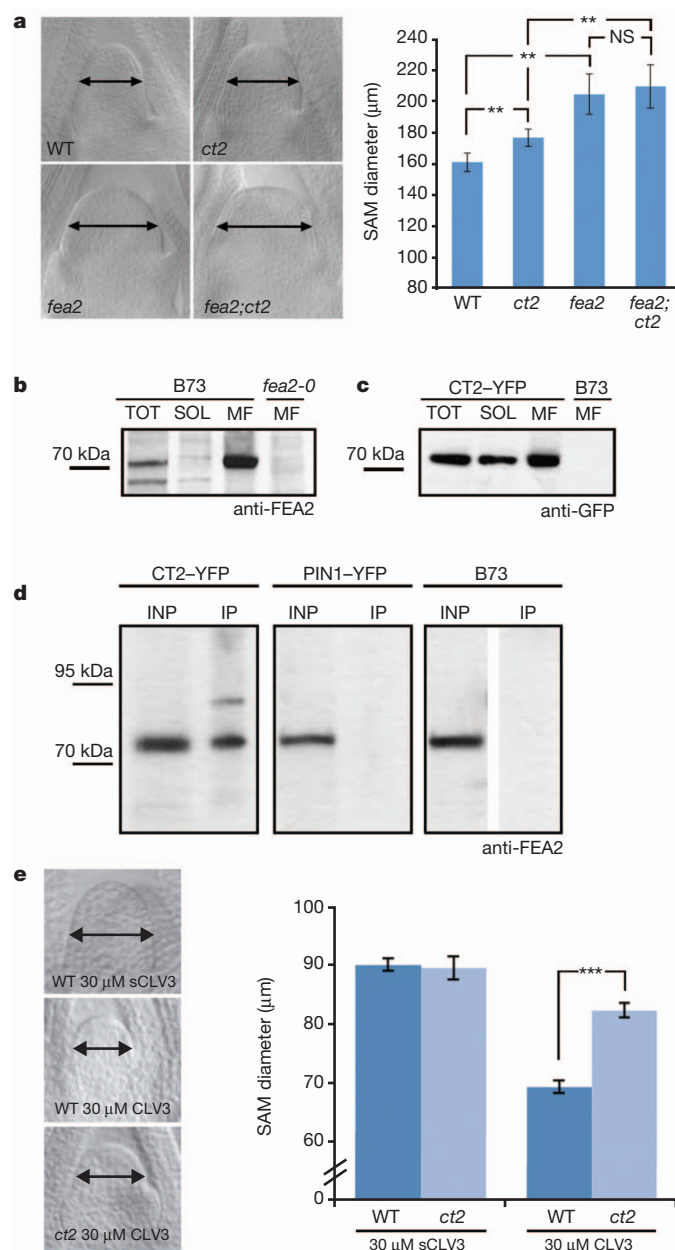


Figure 3 | Interactions between CT2 and CLAVATA signalling. **a**, Cleared SAMs; SAM diameter in double mutants was not significantly different from *fea2-0* (P value = 0.41; Student's t test; n = 10). Error bars represent s.d. NS, not significant; WT, wild type. $**P$ = 0.001. **b**, Western blot showing anti-FEA2 specificity; a band is seen in wild-type (B73) total extracts (TOT) and membrane fractions (MF), but not in soluble (SOL), or *fea2-0*. **c**, CT2-YFP detection using anti-GFP. **d**, CT2 and FEA2 co-immunoprecipitate; a band is detected in co-immunoprecipitates using CT2-YFP, but not in PIN1-YFP or B73 controls. Experiments were conducted with 3 biological replicates. INP, total input. **e**, Embryos cultured with CLV3 or scrambled peptide (sCLV3). Wild-type SAM growth is strongly inhibited by CLV3, but *ct2* is significantly less inhibited ($***P$ value = 0.0001; Student's t test; n = 10 for each genotype). Error bars represent s.e., experiments were conducted with 3 biological replicates.

CLV3-derived signal. Together with our findings that maize *ct2* mutants are strongly fasciated, similar to *CLV* mutants, *ct2* and *fea2* act in a common genetic pathway, and CT2 and FEA2 proteins interact *in vivo*, we suggest that CT2/G α acts to transmit CLAVATA-dependent signals to control shoot stem cell proliferation. This finding helps explain the conundrum that FEA2, like CLV2 and other receptor-like proteins are receptors without a signalling domain. Although it was proposed

that CORYNE (CRN) might signal downstream of CLV2 (ref. 15), this idea has been questioned by the finding that CRN lacks kinase activity⁵.

No obvious meristem size phenotype has been described for G α mutants in other plants, although inducible expression of *Arabidopsis* G α led to production of ectopic shoot meristems¹⁶. A possible explanation for the stronger meristem phenotype in maize is that the ear has undergone recent intense selection for increased size and kernel row number^{17,18} and may be more sensitive to genetic perturbation. Alternatively, the relative importance of G α signalling, or the contribution of parallel pathways with overlapping function(s), may vary between species.

The idea that G α interacts with the single pass transmembrane receptor FEA2 is at odds with the dogma from yeast and mammalian systems that it interacts with G-protein-coupled receptors, which are seven pass transmembrane proteins¹⁹. To ask if this change might be a general phenomenon in plants, we performed mass spectroscopic analysis of proteins immunoprecipitated using CT2-YFP, and found additional predicted LRR receptor proteins as candidate CT2-YFP interactors, and we did not detect any 7TM proteins (Extended Data Table 1). However, based solely on proteomic data it is difficult to rule out the possibility that a 7TM protein acts as an intermediate between FEA2 and CT2. Indeed, plants contain 7TM receptors²⁰, and it is possible that such a protein is in a FEA2-CT2 complex, although their role as GPCRs has been questioned⁷. For example, biochemical and structural data suggests that plant G α proteins are self-activating, supporting the idea that they do not interact with canonical 7TM receptors⁷. In some instances, plant G α proteins are regulated by 7TM regulator of G protein signalling (RGS) proteins. However, such proteins appear to be missing from the grasses, suggesting that G α regulatory mechanisms differ even within the plant kingdom⁷.

Plant G α genes control a wide array of phenotypes, including responses to hormones, drought, pathogens, vegetative growth, flower and panicle development²¹. *Arabidopsis* heterotrimeric G proteins have overlapping functions with the LRR receptor-like kinase *ERECTA* (*ER*), including flower and leaf development and cell division^{22,23}. Double mutant analysis of *er* and *agb1*, which encodes the β -subunit of the heterotrimeric G protein, revealed that *ERECTA* and *AGB1* likely function in the same pathway in regulating fruit shape²². Therefore, as supported by our proteomic data, heterotrimeric G proteins might be involved in signalling downstream of other LRR receptors, which are extremely abundant in plants²⁴.

METHODS SUMMARY

Plant growth and map based cloning. Maize plants were grown in the field or in the greenhouse. Phenotyping used the *ct2-Ref* allele introgressed 5 times into the B73 inbred line. Then 1,000 mutants from a segregating F₂ population were used for map-based cloning, and additional alleles were identified using targeted EMS mutagenesis (Extended Data Fig. 2a). Scanning electron microscopy was performed on fresh tissue using a Hitachi S-3500N SEM, as described¹⁰.

Double mutants were constructed and analysed after genotyping and meristem sizes were measured using cleared tissues.

Transgenic lines and analysis. The CT2-YFP transgene was constructed by amplification of genomic fragments and fusing the YFP gene in-frame at an internal position, and transformed into maize (for primer sequences see Methods). For confocal microscopy, tissues were dissected and counterstained with 1 mg ml⁻¹ FM4-64 solution (Molecular Probes) in water for 1 min then washed with water and imaged within 5 min. For plasmolysis, leaf epidermal tissues were peeled, placed in FM4-64 solution, washed with water twice, and imaged. Subsequently, the tissues were incubated for 5–10 min with 30% glycerol and imaged again. Two-photon images were taken with a custom-made two-photon microscope.

Protein detection and co-immunoprecipitation assays were performed using standard techniques (see Methods).

CLV3 peptide assays. Maize embryos segregating for the *ct2* mutation were dissected at ~10 days after pollination, when the SAM was exposed, and cultured on gel media²⁵ containing CLV3 peptide (RTVPSGPDPLHH; 30 μ g ml⁻¹; Genscript) or scrambled peptide (PPTRGLSHHPVD; 30 μ g ml⁻¹). After 10 days, embryos were harvested for genotyping, and fixed in FAA (formalin, 45%, acetic acid, 10%, ethanol, 45%) and cleared in methyl salicylate, and meristems measured by

microscopy. Experiments used at least 10 embryos per genotype, and were replicated in triplicate.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 January; accepted 20 August 2013.

Published online 11 September 2013.

1. Stahl, Y. & Simon, R. Plant primary meristems: shared functions and regulatory mechanisms. *Curr. Opin. Plant Biol.* **13**, 53–58 (2010).
2. Fletcher, J. C., Brand, U., Running, M. P., Simon, R. & Meyerowitz, E. M. Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. *Science* **283**, 1911–1914 (1999).
3. Clark, S. E., Williams, R. W. & Meyerowitz, E. M. The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* **89**, 575–585 (1997).
4. Jeong, S., Trotochaud, A. E. & Clark, S. E. The *Arabidopsis CLAVATA2* gene encodes a receptor-like protein required for the stability of the *CLAVATA1* receptor-like kinase. *Plant Cell* **11**, 1925–1934 (1999).
5. Nimchuk, Z. L., Tarr, P. T. & Meyerowitz, E. M. An evolutionarily conserved pseudokinase mediates stem cell production in plants. *Plant Cell* **23**, 851–854 (2011).
6. Assmann, S. M. G proteins Go green: a plant G protein signaling FAQ sheet. *Science* **310**, 71–73 (2005).
7. Urano, D. *et al.* G protein activation without a GEF in the plant kingdom. *PLoS Genet.* **8**, e1002756 (2012).
8. Jackson, D., Veit, B. & Hake, S. Expression of maize *KNOTTED1* related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. *Development* **120**, 405–413 (1994).
9. Bommert, P. *et al.* *thick tassel dwarf1* encodes a putative maize ortholog of the *Arabidopsis CLAVATA1* leucine-rich repeat receptor-like kinase. *Development* **132**, 1235–1245 (2005).
10. Taguchi-Shiobara, F., Yuan, Z., Hake, S. & Jackson, D. The *fasciated ear2* gene encodes a leucine-rich repeat receptor-like protein that regulates shoot meristem proliferation in maize. *Genes Dev.* **15**, 2755–2766 (2001).
11. Ashikari, M., Wu, J., Yano, M., Sasaki, T. & Yoshimura, A. Rice gibberellin-insensitive dwarf mutant gene *Dwarf 1* encodes the α -subunit of GTP-binding protein. *Proc. Natl Acad. Sci. USA* **96**, 10284–10289 (1999).
12. van der Hoorn, R. A. *et al.* Structure-function analysis of Cf-9, a receptor-like protein with extracytoplasmic leucine-rich repeats. *Plant Cell* **17**, 1000–1015 (2005).
13. Gallavotti, A., Yang, Y., Schmidt, R. J. & Jackson, D. The relationship between auxin transport and maize branching. *Plant Physiol.* **147**, 1913–1923 (2008).
14. Fiers, M. *et al.* The 14-amino acid CLV3, CLE19, and CLE40 peptides trigger consumption of the root meristem in *Arabidopsis* through a *CLAVATA2*-dependent pathway. *Plant Cell* **17**, 2542–2553 (2005).
15. Müller, R., Bleckmann, A. & Simon, R. The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal *CLAVATA3* independently of *CLAVATA1*. *Plant Cell* **20**, 934–946 (2008).
16. Ullah, H. *et al.* Modulation of cell proliferation by heterotrimeric G protein in *Arabidopsis*. *Science* **292**, 2066–2069 (2001).
17. Bommert, P., Nagasawa, N. S. & Jackson, D. Quantitative variation in maize kernel row number is controlled by the *FASCIATED EAR2* locus. *Nature Genet.* **45**, 334–337 (2013).
18. Brown, P. J. *et al.* Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS Genet.* **7**, e1002383 (2011).
19. Katritch, V., Cherezov, V. & Stevens, R. C. Structure-function of the G protein-coupled receptor superfamily. *Annu. Rev. Pharmacol. Toxicol.* **53**, 531–556 (2013).
20. Pandey, S., Nelson, D. C. & Assmann, S. M. Two novel GPCR-type G proteins are abscisic acid receptors in *Arabidopsis*. *Cell* **136**, 136–148 (2009).
21. Perfus-Barbeoch, L., Jones, A. M. & Assmann, S. M. Plant heterotrimeric G protein function: insights from *Arabidopsis* and rice mutants. *Curr. Opin. Plant Biol.* **7**, 719–731 (2004).
22. Lease, K. A. *et al.* A mutant *Arabidopsis* heterotrimeric G-protein β subunit affects leaf, flower, and fruit development. *Plant Cell* **13**, 2631–2641 (2001).
23. Shpak, E. D., Berthiaume, C. T., Hill, E. J. & Torii, K. U. Synergistic interaction of three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* **131**, 1491–1501 (2004).
24. Shiu, S. H. & Bleeker, A. B. Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci. STKE* **2001**, re22 (2001).
25. Iyer-Pascuzzi, A. S. *et al.* Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant Physiol.* **152**, 1148–1157 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors thank T. Mulligan for plant care; T. Rocheford for growing the *ct2* targeted EMS screens; F. Albeanu for help with two-photon microscopy; P. Yin for assistance with genotyping; K. Rivera and D. Pappin for performing proteomic experiments; and T. Zadrozny, Q. Wu and W. Boss for assistance and advice with two-phase partitioning. P.B. was supported by DFG grant Bo 3012/1-1, A.G. was supported by the BARD Postdoctoral award no FI-431-2008, B.I.J. was supported by Dupont-Pioneer (CSHL-Collaborative agreement), P.B. and D.J. were supported by the National Science Foundation Plant Genome Program, DBI-0604923.

Author Contributions P.B., B.I.J., A.G. and D.J. designed experiments and analysed data; P.B. performed all experiments, except peptide assays and contributions to imaging, phenotyping and biochemical assays (B.I.J.) and fluorescence imaging (A.G.); and P.B. and D.J. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.J. (jacksond@cshl.edu).

Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection

Caroline Goujon¹, Olivier Moncorgé², Hélène Bauby¹, Tomas Doyle¹, Christopher C. Ward¹, Torsten Schaller¹, Stéphane Hué³, Wendy S. Barclay², Reiner Schulz⁴ & Michael H. Malim¹

Animal cells harbour multiple innate effector mechanisms that inhibit virus replication. For the pathogenic retrovirus human immunodeficiency virus type 1 (HIV-1), these include widely expressed restriction factors¹, such as APOBEC3 proteins², TRIM5- α ³, BST2 (refs 4, 5) and SAMHD1 (refs 6, 7), as well as additional factors that are stimulated by type 1 interferon (IFN)^{8–14}. Here we use both ectopic expression and gene-silencing experiments to define the human dynamin-like, IFN-induced myxovirus resistance 2 (MX2, also known as MXB) protein as a potent inhibitor of HIV-1 infection and as a key effector of IFN- α -mediated resistance to HIV-1 infection. MX2 suppresses infection by all HIV-1 strains tested, has equivalent or reduced effects on divergent simian immunodeficiency viruses, and does not inhibit other retroviruses such as murine leukaemia virus. The Capsid region of the viral Gag protein dictates susceptibility to MX2, and the block to infection occurs at a late post-entry step, with both the nuclear accumulation and chromosomal integration of nascent viral complementary DNA suppressed. Finally, human MX1 (also known as MXA), a closely related protein that has long been recognized as a broadly acting inhibitor of RNA and DNA viruses, including the orthomyxovirus influenza A virus^{15,16}, does not affect HIV-1, whereas MX2 is ineffective against influenza virus. MX2 is therefore a cell-autonomous, anti-HIV-1 resistance factor whose purposeful mobilization may represent a new therapeutic approach for the treatment of HIV/AIDS.

We reported previously that IFN- α pre-treatment of cultured human cells and cell lines establishes patterns of HIV-1 inhibition ranging from severe (monocyte-derived macrophages (MDMs), the monocytic line THP-1 and the glioblastoma line U87-MG), to intermediate (primary CD4⁺ T cells), to minimal (lines such as CEM, HUT78 or Jurkat)^{10,17}. We therefore used transcriptional profiling of RNA isolated from 15 cultures (Jurkat, CEM, CEM-SS, HT1080, U87-MG, U937 \pm phorbol 12-myristate 13-acetate (PMA), THP-1 \pm PMA; MDMs from three donors; and CD4⁺ T cells from three donors) in the presence or absence of IFN- α to identify candidate IFN- α -responsive, cell-encoded suppressors of HIV-1 infection (GEO accession number: GSE46599). Two selection criteria were applied to the data: (1) mean IFN- α -mediated induction of >fourfold across all samples; and (2) >fourfold higher expression in MDMs compared to CEM. Fourteen candidate genes were identified (Extended Data Table 1), with *CXCL10*, *STAT1* and *OASL* discounted from further study (the latter being cytotoxic). cDNAs for the remaining 11 genes were inserted into a doxycycline-inducible lentiviral vector, pEasiLV-MCS, in which transgene expression is repressed in vector-producing cells and transduction efficiency of target cells is scored by visualizing expression of E2-Crimson fluorescent protein (Fig. 1a and Methods).

As an initial screen for individual anti-viral capability, parental U87-MG CD4⁺ CXCR4⁺ cultures were untreated or treated with IFN- α , or transduced with high-titre stocks of each vector, as well as with negative control vectors expressing green fluorescent protein (GFP) or CD8, or a positive control expressing the TRIM5-cyclophilin

A (TRIMCyp) fusion protein of owl monkeys, a well-established post-entry inhibitor of HIV-1 (ref. 18). The cultures were induced with doxycycline and >85% of the cells in each culture were confirmed as E2-Crimson-positive (not shown). Five separate wells of each culture were then challenged with one of five escalating doses of HIV-1/Nef-internal ribosome entry signal (IRES)-Renilla, a modified replication-competent virus, and productive infection quantified by monitoring activity of the Renilla luciferase reporter at 48 h (Fig. 1b). Only MX2 exhibited a clear anti-viral phenotype, with the levels of inhibition typically exceeding 90% and approaching those achieved with TRIMCyp or treatment with IFN- α . Similar results were obtained using vesicular stomatitis virus G-glycoprotein (VSV-G)-pseudotyped challenge virus, demonstrating that MX2-mediated inhibition occurs independently of the route of virus entry (Extended Data Fig. 1), as well as with CEM-SS and 293T target cells (Extended Data Fig. 2). The expression profile of MX2 in MDMs, primary T cells and cell lines was assessed by immunoblot (Fig. 1c) and quantitative PCR with reverse transcription (qRT-PCR) (Extended Data Fig. 3), confirming both IFN- α inducibility as well as preferential expression in cells displaying IFN- α -induced resistance to infection^{10,17}.

Having found that ectopic expression of MX2 is sufficient to confer resistance to HIV-1 infection, we used gene silencing to address the contribution of MX2 to the IFN- α -induced anti-viral state. U87-MG CD4⁺ CXCR4⁺ cells were transduced three to four times with either of two lentiviral vectors expressing MX2-specific short hairpin RNAs (shRNAs sh1 and sh2) or a non-targeting shRNA control vector. After at least 8 days, the cultures were incubated with or without IFN- α , challenged with HIV-1/Nef-IRES-Renilla, and infection monitored as Renilla luciferase activity (Fig. 2a). In cultures treated with IFN- α , MX2 silencing stimulated infection by five- to tenfold relative to the control, whereas no effect was noted in the absence of IFN- α , demonstrating that MX2 has a substantial role in the restriction of HIV-1 by IFN- α . Immunoblot analyses confirmed the efficiency of MX2 silencing (Fig. 2b, lanes 4 and 6), and similar results were obtained in a second cell line, THP-1 (Extended Data Fig. 4).

Human MX2 is a member of the IFN-inducible guanosine triphosphatase (GTPase) superfamily that includes proteins involved in cellular processes requiring membrane remodelling, such as vesicular transport and cytokinesis, as well as in resistance to intracellular pathogens¹⁹. The most closely related family member is human MX1 (63% amino acid sequence identity), which inhibits a variety of RNA and DNA viruses, including influenza A virus, La Crosse encephalitis virus and hepatitis B virus, and is thought to form an oligomeric ring that engages and disrupts viral nucleoprotein/replication complexes^{15,20,21}. Conversely, relatively little information concerning MX2 function is available: it is nuclear as well as cytoplasmic and accumulates at the cytoplasmic face of nuclear pore complexes. MX2 may have a role in cell cycle progression, but has not previously been ascribed notable anti-viral function^{14,16,22,23}.

To define more closely how MX2 inhibits HIV-1 replication, we challenged parental U87-MG CD4⁺ CXCR4⁺ cells, cultured with or

¹Department of Infectious Diseases, King's College London, London SE1 9RT, UK. ²Section of Virology, Division of Infectious Diseases, Imperial College London, London W2 1PG, UK. ³Centre for Medical Molecular Virology, Division of Infection and Immunity, University College London, London WC1 6RT, UK. ⁴Department of Medical and Molecular Genetics, King's College London, London SE1 9RT, UK.

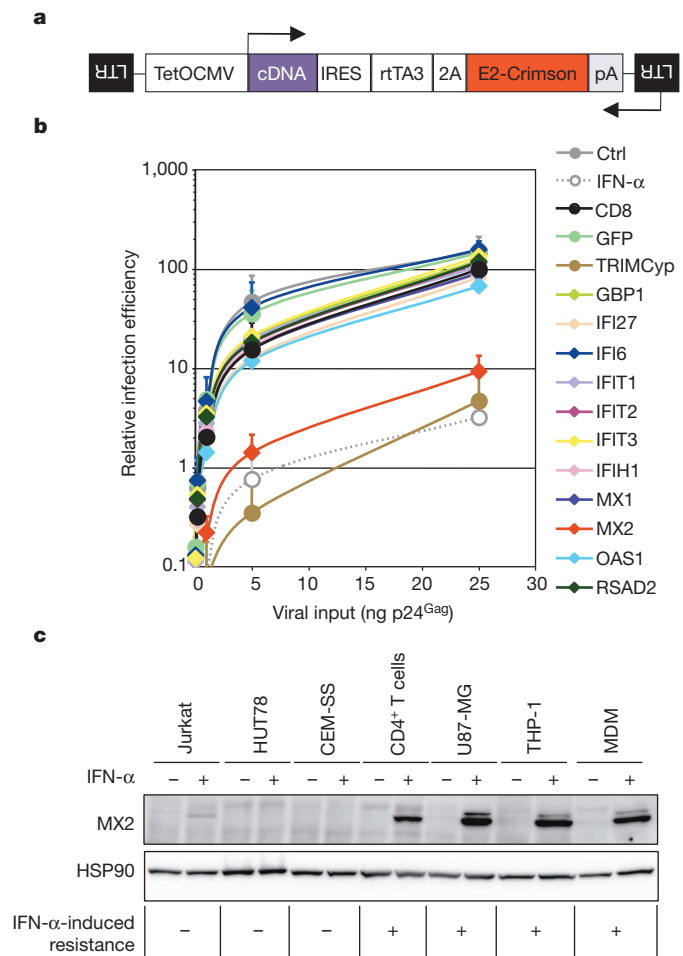


Figure 1 | Human MX2 is a potent inhibitor of HIV-1 infection.

a, Schematic representation of the EasiLV (E2-Crimson antisense inducible lentiviral vector) system. pEasiLV-MCS contains an internal antisense and Tet-inducible expression cassette driving expression of a tricistronic RNA encoding the cDNA of interest, the reverse responsive tetracycline transactivator variant 3 (rtTA3) and the E2-Crimson indicator gene (TetOCMV, tetracycline operator-minimal cytomegalovirus promoter; pA, polyadenylation signal) (see Methods for details). **b**, Candidate cDNA screen in U87-MG CD4⁺ CXCR4⁺ cells. U87-MG CD4⁺ CXCR4⁺ cells were transduced with EasiLV expressing the different candidate cDNAs, CD8 (negative control), GFP (negative control) or TRIMCyp (positive control) cDNAs and either treated with doxycycline for 48 h, left untransduced (Ctrl) or treated with 1,000 U ml⁻¹ IFN- α for 24 h before HIV-1 infection. The cells were infected with increasing viral inputs of NL4-3/Nef-IRES-Renilla (0.04–25 ng p24^{Gag}) and infection efficiency was monitored 48 h later by measuring Renilla activity. Mean relative infection efficiencies with standard deviations from four independent experiments are shown. **c**, Immunoblot analysis of MX2 protein levels in control and IFN- α -treated Jurkat, HUT78, CEM-SS, primary CD4⁺ T cells, U87-MG, THP-1 and MDMs; HSP90 served as a loading control. The IFN- α -induced resistance phenotype of each cell type is shown underneath (–, no resistance; +, resistance).

without IFN- α , and cells transduced with CD8- or MX2-expressing vectors, with wild-type HIV-1 and then collected total DNA at 2, 6, 24 and 48 h. The 48-h cultures were also analysed for p24^{Gag} expression using flow cytometry, confirming MX2-mediated inhibition of viral gene expression (Extended Data Fig. 5). qPCR was then used to measure viral reverse transcripts representing three phases of replication: extended minus (first)-strand cDNA, 2-long terminal repeat (LTR) circular DNA (a marker for viral cDNA nuclear localization) and integrated (provirus) DNA (Fig. 3). As reported previously, IFN- α treatment severely blocked the accumulation of all HIV-1 cDNAs¹⁰. By contrast, MX2 did not measurably affect the synthesis or accumulation of minus-strand cDNA, but reduced the levels of 2-LTR circles

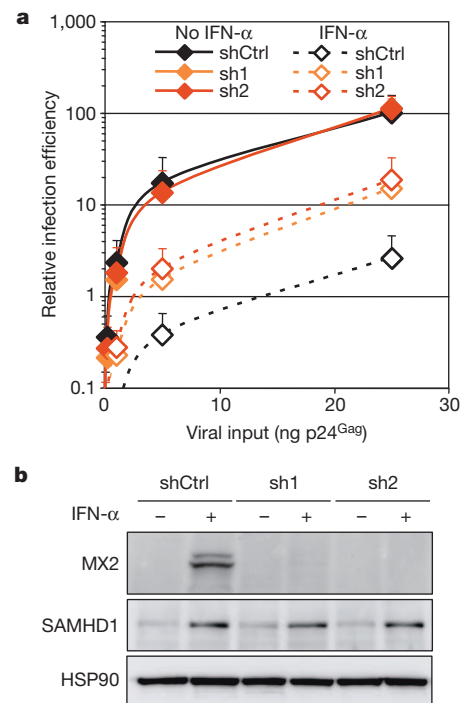


Figure 2 | MX2 is required for effective IFN- α -induced suppression of HIV-1. **a**, U87-MG CD4⁺ CXCR4⁺ cells expressing a control shRNA or two different shRNAs targeting MX2 were cultured with or without IFN- α (500 U ml⁻¹) for 24 h. Cells were infected with five different doses of NL4-3/Nef-IRES-Renilla (0.04–25 ng p24^{Gag}) for 48 h, and Renilla activity was measured. Mean relative infection efficiencies from two independent experiments are shown. **b**, Immunoblot analysis of parallel samples from **a**. Protein levels of MX2 and SAMHD1 (positive control for IFN- α induction) were determined and HSP90 served as a loading control.

and proviruses by ~90%, possibly indicating a blockade to the nuclear uptake of viral replication complexes or a decrease in their stability.

We next examined the ability of MX2 to suppress infection by a range of primate lentiviruses including laboratory-adapted strains of HIV-1, HIV-1-transmitted founder strains, HIV-2 and simian immunodeficiency viruses (SIVs) derived from the rhesus macaque (SIV_{MAC}), mandrill (SIV_{MND}) or African green monkey (SIV_{AGM}). This was quantified by using the virus-encoded Tat proteins that are expressed after infection to trans-activate an HIV-LTR/luciferase reporter cassette that was resident in target U87-MG cells. These reporter cells were transduced with either CD8- or MX2-expressing vectors and subsequently challenged with two doses of VSV-G-pseudotyped stocks of HIVs or SIVs. Measurement of luciferase levels at 48 h showed that all HIV-1s and SIV_{MND} were susceptible to potent repression by MX2, whereas HIV-2, SIV_{MAC} and SIV_{AGM} were somewhat less sensitive (Fig. 4a). The analysis was then extended to three non-primate viruses, the lentiviruses equine infectious anaemia virus (EIAV) and feline immunodeficiency virus (FIV), and the gammaretrovirus murine leukaemia virus (MLV). Here, we used retroviral vectors encoding GFP and monitored single-cycle infectivity at 48 h by flow cytometry (Fig. 4b). Interestingly, whereas MX2 suppressed infection by the HIV-1-based vector by ~80%, no inhibition of the three non-primate viruses was observed, demonstrating that the human MX2 protein exhibits substrate selectivity, albeit to differing extents, for primate lentiviruses.

Current views on the post-entry progression of HIV-1 infection invoke the sustained presence of the viral Capsid (CA) protein within reverse transcription complexes, as well as a central role for CA in mediating interactions with host proteins such as cyclophilin A, TNPO3, NUP358 (also known as RANBP2), NUP153 or TRIM5- α that influence the fate of infection^{1,3,24–27}. To address whether CA

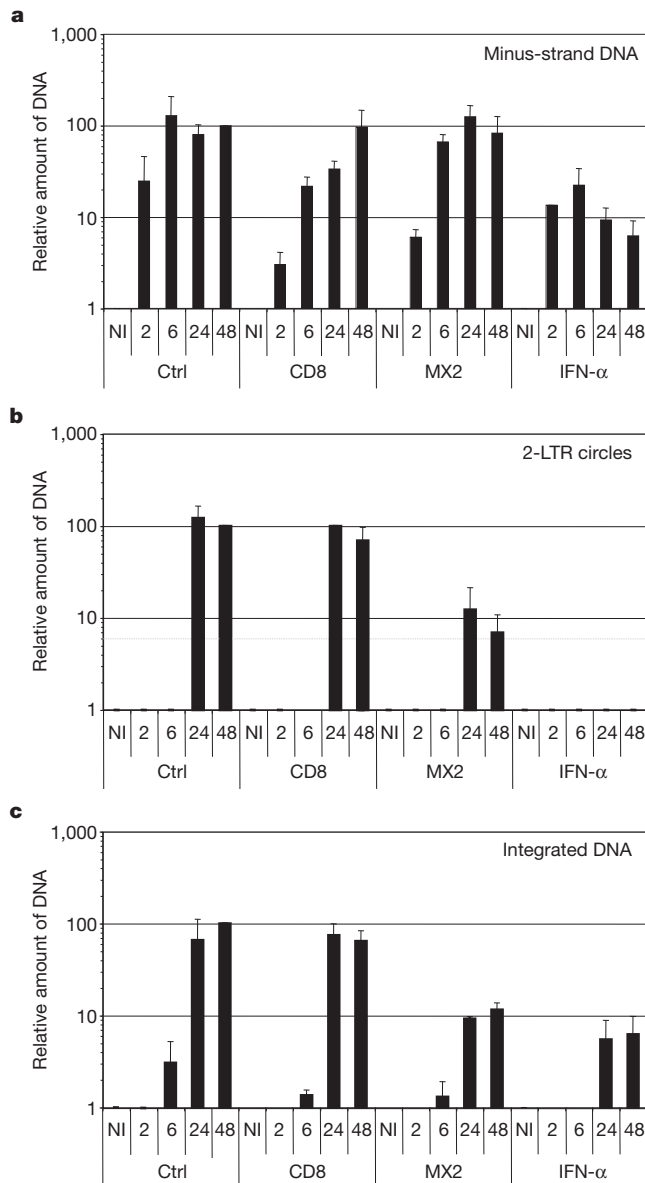


Figure 3 | MX2 inhibits the nuclear accumulation and integration of HIV-1 reverse transcripts. **a–c**, U87-MG CD4⁺ CXCR4⁺ cells were transduced with EasiLV expressing CD8 or MX2 and treated with doxycycline for 48 h, left untransduced (Ctrl) or treated with IFN- α for 24 h before infection. The cells were either not infected (NI) or challenged with 10 ng p24^{Gag} HIV-1_{IIIB} and collected at 2, 6, 24 or 48 h after infection for DNA extraction and qPCR analysis of minus-strand DNA (**a**), 2-LTR circle DNA (**b**) and integrated proviral DNA (**c**). Mean values of relative amounts of DNA (normalized to control at 48 h) from three independent experiments are shown. The detection limit for 2-LTR circle qPCR was ten copies per reaction, which corresponds to ~6% relative copies as indicated on the graph by a dashed grey line. p24^{Gag} expression was also determined at 48 h in parallel samples to monitor productive infection (Extended Data Fig. 5).

determines the sensitivity of HIV-1 to MX2, we measured the effects of MX2 using GFP-encoding vectors carrying the P90A or N74D mutations in CA that inhibit/prevent interactions with CypA, TNPO3, NUP358 or NUP153, or that had the CA region of Gag replaced with SIV_{MAC} CA (Fig. 4c). In contrast to the ~80% inhibition of wild-type Gag, the P90A and N74D CA variants were insensitive or only mildly sensitive to inhibition by MX2, respectively, and the SIV-CA-containing chimera displayed modest inhibition, reflecting closely that of the parental SIV_{MAC} protein. The observation that modifying HIV-1 CA can control MX2 susceptibility or escape suggests that CA is a specific target of MX2.

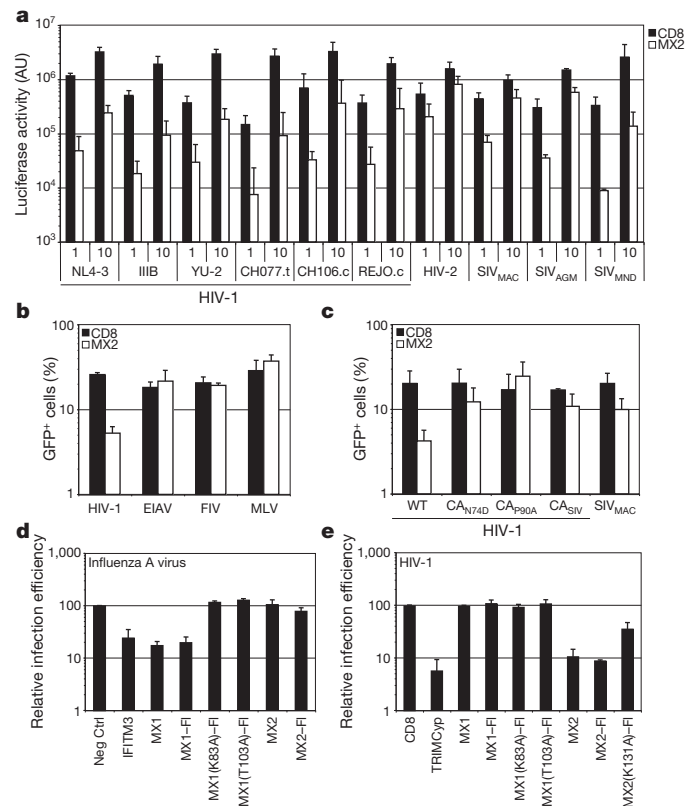


Figure 4 | Viral substrates for the human MX1 and MX2 proteins. **a**, U87-MG/LTR-Luc cells were transduced with EasiLV expressing CD8 or MX2. Cells were infected with two doses (1 and 10, corresponding to 50 and 500 pg RT) of VSV-G-pseudotyped HIV-1_{NL4-3}, HIV-1_{IIIB}, HIV-1_{YU-2}, HIV-1_{CH077.1}, HIV-1_{CH106.c}, HIV-1_{REJO.c}, HIV-2_{ROD10}, SIV_{MAC239}, SIV_{AGMTAN} and SIV_{MND121}. Luciferase activity was measured at 48 h. Mean values for three independent experiments are shown. **b**, CD8- or MX2-expressing U87-MG cells were challenged with HIV-1, ELAV-, FIV- and MLV-based retroviral vectors expressing GFP at a multiplicity of infection (m.o.i.) of 0.25. The percentage of GFP-expressing cells was evaluated by flow cytometry. Mean percentages of transduced cells from four independent experiments are shown. **c**, CD8- or MX2-expressing U87-MG cells were challenged with GFP-encoding HIV-1-based vectors (containing wild-type (WT) CA, CA_{N74D}, CA_{P90A} or CA from SIV_{MAC} (CA_{SIV})), or an SIV_{MAC}-based vector at a m.o.i. of 0.25 as in **b**. The percentage of GFP-expressing cells was evaluated, and mean percentages of transduced cells for four independent experiments (three for CA_{SIV}) are shown. **d**, 293T cells were co-transfected with expression plasmids for GFP (Neg Ctrl), IFITM3, untagged and Flag-tagged MX1 and MX2 (MX1-FI and MX2-FI), or the Flag-tagged MX1 GTPase-deficient mutants MX1(K83A) and MX1(T103A) along with an influenza A virus firefly luciferase minigenome plasmid and a Renilla luciferase expression plasmid. At 24 h, cells were infected with influenza A virus A/Victoria/3/75 (H3N2) at a m.o.i. of 2 and firefly and Renilla luciferase activities were measured 18 h after infection. Mean relative infection efficiencies for three independent experiments are shown. **e**, U87-MG CD4⁺ CXCR4⁺ cells were transduced with EasiLV expressing CD8, TRIMCyp, MX1, MX2 or the mutants MX1(K83A), MX1(T103A) and MX2(K131A). The cells were infected with 25 ng p24^{Gag} of NL4-3/Nef-IRES-Renilla and infection efficiency was monitored at 48 h by measuring Renilla activity. Mean relative infection efficiencies from three independent experiments are shown.

In a final series of experiments we assessed the effects of MX1 and MX2 on HIV-1 and influenza A virus replication (using analogous assays that measure the culmination of infection, viral RNA synthesis and protein expression). Influenza A virus genome segment replication was determined by co-transfecting 293T cells with a vector expressing a firefly luciferase-containing minigenome (as well as a vector expressing Renilla luciferase for normalization), together with vectors for wild-type MX1 or MX2 (Flag-tagged and -untagged), or the tagged GTPase-deficient MX1 derivatives K83A and T103A^{28,29}. At 24 h, the

cultures were infected with influenza A virus, and firefly luciferase expression measured 18 h later (Fig. 4d). As established previously, wild-type MX1, as well as IFITM3 (positive control)³⁰, suppressed replication by 75–80%, whereas the GTPase domain mutant proteins had lost anti-viral activity^{28,29}; consistent with previous studies, MX2 did not exert any inhibitory effect¹⁶. The wild-type MX1 and MX2 proteins, and the K131A mutant of MX2 that does not bind GTP²², were then examined for their HIV-1 inhibitory phenotypes in transduced U87-MG CD4⁺ CXCR4⁺ cells as in Fig. 1b. In contrast to the results with influenza A virus, MX1 had no effect on HIV-1, and the mutated MX2(K131A) protein still retained a degree of anti-viral function (~65% inhibition, Fig. 4e). Immunoblotting confirmed expression of the Flag-tagged proteins, although MX2(K131A) accumulated to a lower level than the wild-type protein (Extended Data Fig. 6).

Here we describe the identification of human MX2 as an IFN- α -inducible anti-retroviral effector that, among primate immunodeficiency viruses, is most potent against HIV-1, but does not affect the non-primate viruses MLV, EIAV and FIV (Fig. 4a, b). Understanding the molecular details of MX2's recognition and inactivation of post-entry viral reverse transcription complexes, the interplay with other regulatory host proteins that interact with CA, and the basis for the dichotomy between MX2/HIV-1 inhibition and MX1/influenza virus inhibition with respect to GTPase function (Fig. 4d, e) will help to elucidate the mechanism of this new mode of cell-mediated resistance to retroviral infection. As viral inhibition occurs relatively late during infection and is manifested as the failure to accumulate viral cDNA in the nucleus (Fig. 3), the anti-viral action of MX2 is distinct from TRIM5- α - or APOBEC3G-mediated inhibition of early reverse transcription or SAMHD1-mediated restriction through deoxynucleotide triphosphate depletion¹.

Last, we note that although MX2 silencing substantially relieves IFN- α -induced resistance to HIV-1, measurable inhibition persists (Fig. 2 and Extended Data Fig. 4); taken together with the observation that IFN- α imposes an early block to HIV-1 reverse transcription (Fig. 3)¹⁰, we speculate that additional IFN-stimulated factor(s) that interfere with the initial post-entry phases of HIV-1 infection remain to be discovered.

METHODS SUMMARY

Plasmids, cells, viral vectors and EasiLV system. All reagents, including the novel inducible lentivirus vector pEasiLV-MCS, are described in Methods. Candidate cDNAs were cloned into pEasiLV-MCS for functional screening.

Virus infection. Lentiviral, retroviral and influenza A virus infections were monitored using standard reporter genes, and HIV-1 cDNA was measured by qPCR.

Microarray. Illumina HumanHT12v4 expression bead chips were probed with RNA from primary cells and cell lines, treated or not treated with IFN- α .

MX2 silencing. MX2 silencing was achieved with shRNAs expressed from a lentiviral vector, generated using a modified version of pAPM, and primer sequences available on the Open Biosystems website (<http://www.thermoscientificbio.com/rnai-and-custom-rna-synthesis/shrna/gipz-lentiviral-shrna/>) (target sequences: MX2-1, 5'-AAGATGTCCTTTCTAATG-3'; MX2-2, 5'-CCAACCAGATCCC ATTTAT-3').

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 June; accepted 14 August 2013.

Published online 18 September 2013.

- Malim, M. H. & Bieniasz, P. D. HIV restriction factors and mechanisms of evasion. *Cold Spring Harb. Perspect. Med.* **2**, a006940 (2012).
- Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).
- Stremlau, M. *et al.* The cytoplasmic body component TRIM5 α restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853 (2004).
- Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).
- Van Damme, N. *et al.* The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* **3**, 245–252 (2008).

- Hrecka, K. *et al.* Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* **474**, 658–661 (2011).
- Laguet, N. *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* **474**, 654–657 (2011).
- Asmuth, D. M. *et al.* Safety, tolerability, and mechanisms of antiretroviral activity of pegylated interferon Alfa-2a in HIV-1-monoinfected participants: a phase II clinical trial. *J. Infect. Dis.* **201**, 1686–1696 (2010).
- Bitzegeio, J., Sampias, M., Bieniasz, P. D. & Hatzioannou, T. Adaptation to the interferon-induced antiviral state by human and simian immunodeficiency viruses. *J. Virol.* **87**, 3549–3560 (2013).
- Goujon, C. & Malim, M. H. Characterization of the alpha interferon-induced postentry block to HIV-1 infection in primary human macrophages and T cells. *J. Virol.* **84**, 9254–9266 (2010).
- Ho, D. D. *et al.* Recombinant human interferon alfa-A suppresses HTLV-III replication *in vitro*. *Lancet* **325**, 602–604 (1985).
- Meylan, P. R., Guatelli, J. C., Munis, J. R., Richman, D. D. & Kornbluth, R. S. Mechanisms for the inhibition of HIV replication by interferons- α , - β , and - γ in primary human macrophages. *Virology* **193**, 138–148 (1993).
- Pillai, S. K. *et al.* Role of retroviral restriction factors in the interferon- α -mediated suppression of HIV-1 *in vivo*. *Proc. Natl Acad. Sci. USA* **109**, 3035–3040 (2012).
- Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485 (2011).
- Haller, O. & Kochs, G. Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity. *J. Interferon Cytokine Res.* **31**, 79–87 (2011).
- Pavlovic, J., Zurcher, T., Haller, O. & Staeheli, P. Resistance to influenza virus and vesicular stomatitis virus conferred by expression of human MxA protein. *J. Virol.* **64**, 3370–3375 (1990).
- Goujon, C. *et al.* Evidence for IFN α -induced, SAMHD1-independent inhibitors of early HIV-1 infection. *Retrovirology* **10**, 23 (2013).
- Sayah, D. M., Sokolskaja, E., Berthou, L. & Luban, J. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* **430**, 569–573 (2004).
- Kim, B. H., Shenoy, A. R., Kumar, P., Bradfield, C. J. & MacMicking, J. D. IFN-inducible GTPases in host cell defense. *Cell Host Microbe* **12**, 432–444 (2012).
- Gao, S. *et al.* Structure of myxovirus resistance protein A reveals intra- and intermolecular domain interactions required for the antiviral function. *Immunity* **35**, 514–525 (2011).
- Mitchell, P. S. *et al.* Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. *Cell Host Microbe* **12**, 598–604 (2012).
- King, M. C., Raposo, G. & Lemmon, M. A. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. *Proc. Natl Acad. Sci. USA* **101**, 8957–8962 (2004).
- Melén, K. *et al.* Human MxB protein, an interferon- α -inducible GTPase, contains a nuclear targeting signal and is localized in the heterochromatin region beneath the nuclear envelope. *J. Biol. Chem.* **271**, 23478–23486 (1996).
- Brass, A. L. *et al.* Identification of host proteins required for HIV infection through a functional genomic screen. *Science* **319**, 921–926 (2008).
- König, R. *et al.* Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* **135**, 49–60 (2008).
- Lee, K. *et al.* Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).
- Shah, V. B. *et al.* The host proteins transportin SR2/TNPO3 and cyclophilin A exert opposing effects on HIV-1 uncoating. *J. Virol.* **87**, 422–432 (2013).
- Pitossi, F. *et al.* A functional GTP-binding motif is necessary for antiviral activity of Mx proteins. *J. Virol.* **67**, 6726–6732 (1993).
- Ponten, A., Sick, C., Weeber, M., Haller, O. & Kochs, G. Dominant-negative mutants of human MxA protein: domains in the carboxy-terminal moiety are important for oligomerization and antiviral activity. *J. Virol.* **71**, 2591–2599 (1997).
- Brass, A. L. *et al.* The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell* **139**, 1243–1254 (2009).

Acknowledgements We wish to thank L. Apolonia, A. Cimarelli, B. Hahn, T. Hatzioannou, J. Kappes, P. Mangeot, S. Papaioannou, N. Parrish, N. Sherer and C. Swanson for the provision of reagents and for discussions, and M. Mirza, E. Papouli, Q. Oscar Y. Li and G. Pacini for technical assistance. This work was supported by the UK Medical Research Council, the National Institutes of Health (DA033773), the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. PIEF-GA-2009-237501 (to C.G.), a Wellcome Trust Research Training Fellowship (to T.D.) and the Department of Health via a National Institute for Health Research comprehensive Biomedical Research Centre award to Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust.

Author Contributions C.G. and M.H.M. designed the study and wrote the manuscript. C.G. carried out the experiments. O.M. and W.S.B. designed and carried out the influenza A virus experiment. O.M., H.B., T.D., C.C.W., T.S. and S.H. provided technical assistance. R.S. performed the microarray analysis. All authors read and approved the final manuscript.

Author Information The microarray methods and data are available under Gene Expression Omnibus (GEO) accession number GSE46599. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.H.M. (michael.malim@kcl.ac.uk).

MX2 is an interferon-induced inhibitor of HIV-1 infection

Melissa Kane^{1,2}, Shalini S. Yadav^{1,2,3†}, Julia Bitzegeio^{1,2}, Sebla B. Kutluay^{1,2}, Trinity Zang^{1,2,3}, Sam J. Wilson^{1,2,3†}, John W. Schoggins^{4†}, Charles M. Rice⁴, Masahiro Yamashita¹, Theodora Hatzioannou¹ & Paul D. Bieniasz^{1,2,3}

HIV-1 replication can be inhibited by type I interferon (IFN), and the expression of a number of gene products with anti-HIV-1 activity is induced by type I IFN^{1,2}. However, none of the known antiretroviral proteins can account for the ability of type I IFN to inhibit early, preintegration phases of the HIV-1 replication cycle in human cells^{3,4}. Here, by comparing gene expression profiles in cell lines that differ in their ability to support the inhibitory action of IFN- α at early steps of the HIV-1 replication cycle, we identify myxovirus resistance 2 (MX2) as an interferon-induced inhibitor of HIV-1 infection. Expression of MX2 reduces permissiveness to a variety of lentiviruses, whereas depletion of MX2 using RNA interference reduces the anti-HIV-1 potency of IFN- α . HIV-1 reverse transcription proceeds normally in MX2-expressing cells, but 2-long terminal repeat circular forms of HIV-1 DNA are less abundant, suggesting that MX2 inhibits HIV-1 nuclear import, or destabilizes nuclear HIV-1 DNA. Consistent with this notion, mutations in the HIV-1 capsid protein that are known, or suspected, to alter the nuclear import pathways used by HIV-1 confer resistance to MX2, whereas preventing cell division increases MX2 potency. Overall, these findings indicate that MX2 is an effector of the anti-HIV-1 activity of type-I IFN, and suggest that MX2 inhibits HIV-1 infection by inhibiting capsid-dependent nuclear import of subviral complexes.

We and others have previously identified proteins with antiretroviral activity on the basis of their differential expression in cells that are permissive or non-permissive with respect to particular steps in the HIV-1 life cycle^{5,6}. We noticed that monocytoid cell lines varied in their ability to support the anti-HIV-1 activity of type I IFN. Specifically, IFN- α treatment of THP-1 cells caused an ~40-fold reduction in infection by an HIV-1-based green fluorescent protein (GFP) reporter vector, whereas treatment of K562 and U937 cells had little effect (Fig. 1a). When these cell lines were differentiated into a macrophage-like state by treatment with phorbol 12-myristate 13-acetate (PMA), the inhibitory effect of IFN- α was accentuated in THP-1 cells and accentuated to a lesser extent in U937 cells, but remained nearly absent in K562 cells (Fig. 1a).

To identify candidate effectors of the antiviral action of IFN- α , we used microarrays to measure messenger RNA levels in the aforementioned cell lines. Twenty-two genes whose induction, or non-induction, by IFN- α correlated to varying degrees with the ability or inability of IFN- α to inhibit HIV-1–GFP vector infection in the monocytoid cell lines were selected for further study (Fig. 1b and Extended Data Figs 1 and 2). Among these candidates, MX2, a gene that was not previously thought to exhibit antiviral activity⁷, was of particular interest as we recently identified it as a ‘hit’ in an overexpression screen in a T-cell line during which MX2 modestly inhibited infection by HIV-1 (ref. 8). Western blot analyses confirmed that MX2 expression was strongly induced by IFN- α in THP-1 cells but not K562 cells, and a basal level of MX2 expression was slightly increased by IFN- α treatment in U937 cells

(Fig. 1c). MX2 was expressed at a basal level in primary CD4⁺ T cells and macrophages, and was induced to varying degrees by IFN- α , depending on the individual donor, and how cells were activated (Extended Data Fig. 3).

Expression of the 22 candidate and control genes in K562 cells revealed that only MX2 and a control antiviral gene coding for rhesus macaque TRIM5- α ⁹ inhibited HIV-1 infection. (Fig. 2a). A rhesus macaque variant of MX2 also inhibited HIV-1 infection to a similar degree as human MX2, whereas MX1 was inactive against HIV-1 (Fig. 2a), even though it inhibits a variety of other viruses⁷. Although MX2 clearly inhibited HIV-1 infection (Fig. 2a–d), the fact that U937 cells (Fig. 1a), primary macrophages and anti-CD3/CD28-stimulated CD4⁺ T cells are readily infected by HIV-1, despite expressing appreciable levels of MX2 (Fig. 1c and Extended Data Fig. 3), indicates that the block imposed by MX2 is not absolute, or that MX2 potency is perhaps influenced by the cellular environment or cofactors.

MX1 and MX2 are members of a family of dynamin-like GTPases⁷, but only MX2 is localized to the nucleus by virtue of a basic nuclear localization signal (NLS) contained within its amino-terminal 25 amino acids^{10,11}. Notably, the N-terminal 25 amino acids that encode the MX2 NLS were strictly required for antiviral activity (Fig. 2b, c). Conversely, the mutations K131A and T151A—which inhibit GTP binding and hydrolysis, respectively¹¹—did not block the anti-HIV-1 activity of MX2 (Fig. 2b, c). This result is in contrast to findings with MX1, whose antiviral activity is GTPase dependent⁷, but should be interpreted cautiously given the reported ability of these MX2 mutants to induce a generalized perturbation of nucleocytoplasmic transport¹¹. In addition to its activity against HIV-1 and HIV-2 (Fig. 2d), MX2 expression in HOS cells inhibited infection by GFP reporter viruses based on a variety of primate lentiviruses, including simian immunodeficiency viruses SIV_{MAC}, SIV_{AGM}Tan and SIV_{AGM}Sab, with some variation in MX2 antiviral potency (Fig. 2e). The nonprimate lentiviruses—equine infectious anaemia virus and feline immunodeficiency virus—were less potently inhibited, whereas a gammaretrovirus—murine leukaemia virus—was only marginally sensitive to MX2.

The experiments described above all represented single-cycle infection assays, using vesicular stomatitis virus glycoprotein (VSV-G)-pseudotyped reporter viruses. However, expression of MX2 in GHOST-R5 cells also inhibited infection by two full-length primary HIV-1 strains, suggesting that MX2 inhibition was independent of the route of entry, and not counteracted by HIV-1 accessory genes (Fig. 3a). Moreover, MX2 expression in GHOST-X4 cells inhibited spreading infection by full-length replication-competent HIV-1_{NL4-3} (Fig. 3b), reducing the number of infected cells by ~20-fold during the exponential phase of viral growth. Reduction of MX2 expression in THP-1 cells (Fig. 3c, d) or in HOS cells (Extended Data Fig. 4a, b) using short hairpin RNAs (shRNAs) reduced, but did not eliminate, the antiviral effect of IFN- α . Thus, MX2 is required for the full potency of IFN- α ,

¹Aaron Diamond AIDS Research Center, New York, New York 10016, USA. ²Laboratory of Retrovirology, The Rockefeller University, New York, New York 10065, USA. ³Howard Hughes Medical Institute, New York, New York 10016, USA. ⁴Center for the Study of Hepatitis C, The Rockefeller University, New York, New York 10065, USA. [†]Present addresses: University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9048, USA (J.W.S.); MRC Centre for Virus Research, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK (S.J.W.); Weill Cornell Medical College, 525 East 68th street, New York, New York 10025, USA (S.S.Y.).

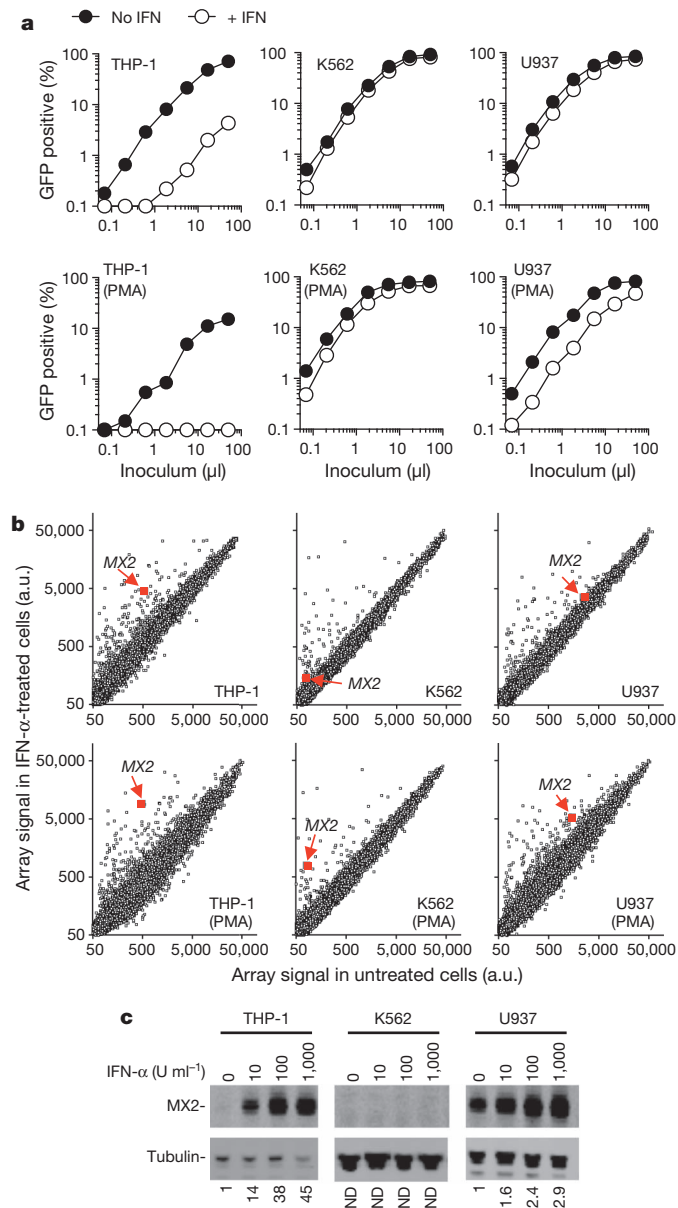


Figure 1 | Differential effects of IFN- α on HIV-1 infection of monocytoid cell lines correlates with MX2 expression. **a**, Undifferentiated (top) or PMA-differentiated (bottom) THP-1, K562 and U937 cells with or without IFN- α treatment ($1,000\ U\ ml^{-1}$) were challenged with a GFP-expressing HIV-1 vector (CSGW). **b**, RNA extracted from cells treated identically to those shown in **a** was analysed on microarrays. The array signal is plotted in arbitrary units (a.u.), and the data points representing MX2 are highlighted. **c**, Western blot analysis of MX2 and tubulin expression in monocytoid cell lines treated for 24 h with the indicated doses of IFN- α . Numbers below each lane indicated fold increase in MX2 protein levels relative to untreated cells. ND, not detected.

but is not solely responsible for the inhibitory action of IFN- α on the early steps of the HIV-1 replication cycle.

Consistent with this conclusion, IFN- α treatment reduced the accumulation of HIV-1 reverse transcripts in HOS cells (Fig. 4a), as has previously been reported for other cell types¹². Conversely, MX2 expression did not inhibit reverse transcript accumulation in either HOS or K562 cells (Fig. 4a and Extended Data Fig. 5). However, MX2 did reduce the generation of 2-long terminal repeat (2-LTR) circles (Fig. 4a and Extended Data Fig. 5), which are thought to form only after retroviral DNA has accessed the nucleus of infected cells. MX2 may, therefore, inhibit the entry of HIV-1 into the nucleus, or perhaps cause destabilization of viral DNA in the nucleus. Consistent with previous

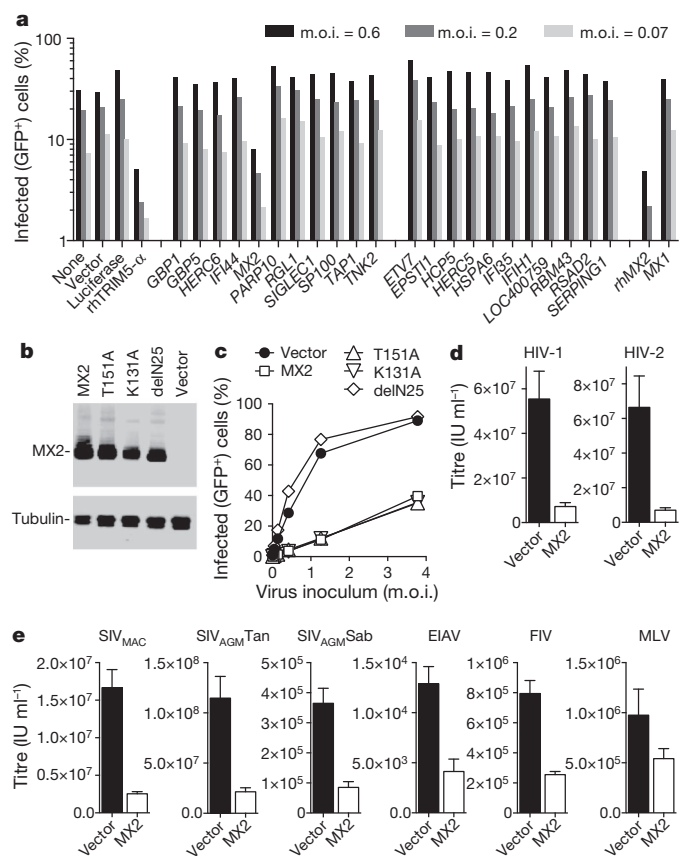


Figure 2 | Inhibition of lentivirus infection by wild-type and mutant MX2, but not other differentially interferon-induced genes. **a**, Infection of K562 cells, previously transduced with an HIV-1 vector (SCRPSY) expressing negative (Luciferase) or positive (rhesus macaque (rh)TRIM5- α -coding) control genes, or candidate antiviral genes, with GFP-expressing HIV-1 vector (CSGW) at the indicated multiplicity of infection (m.o.i.). **b**, Western blot analysis of MX2 and tubulin expression in K562 cell clones transduced with an HIV-1 vector (CSIB) expressing wild-type and mutant MX2 proteins. delN25, MX2 mutant lacking the N-terminal NLS. **c**, Infection of the same K562 cells as in **b** with an HIV-1-GFP reporter virus. **d**, **e**, Infection of HOS cells, previously transduced with an MX2-expressing or empty HIV-1 vector (SCRPSY), with various GFP reporter viruses. Titres are mean \pm s.d., $n = 3$ technical replicates, representative of four experiments. EIAV, equine infectious anaemia virus; FIV, feline immunodeficiency virus; MLV, murine leukaemia virus.

reports^{10,11}, we found that that N- or carboxy-terminally haemagglutinin-tagged forms of MX2 were particularly concentrated at nuclear pores marked by the nucleoporin NUP98 (Extended Data Fig. 6). The MX2(K131A) mutant is primarily cytoplasmic but nevertheless inhibits nucleocytoplasmic transport¹¹ and also retains antiviral activity (Fig. 2c). Therefore, alteration of the fate of incoming HIV-1 DNA with respect to the nucleus may underlie the antiviral activity of MX2, even though stable physical association with nuclear pores may not be required for antiviral function.

The HIV-1 capsid protein (CA) is a key determinant required for infection of non-dividing cells and nuclear entry of subviral complexes^{13–15}. Indeed, HIV-1 CA mutations have been shown to change the requirement for specific nucleoporins (for example, NUP358 (also known as RANBP2), NUP85, NUP153, NUP155) during HIV-1 infection, and to alter the distribution of sites at which HIV-1 DNA integrates into host chromosomes^{16–18}. Therefore, we tested whether a number of CA mutations that are known or suspected to affect the pathway used by HIV-1 DNA into the nucleus also affected sensitivity to inhibition by MX2 (Fig. 4b). Of these, a mutation (N57S) that confers cell cycle dependence on HIV-1 infection¹⁹, and presumably restricts HIV-1 nuclear entry to the mitotic phase of the cell cycle, conferred resistance

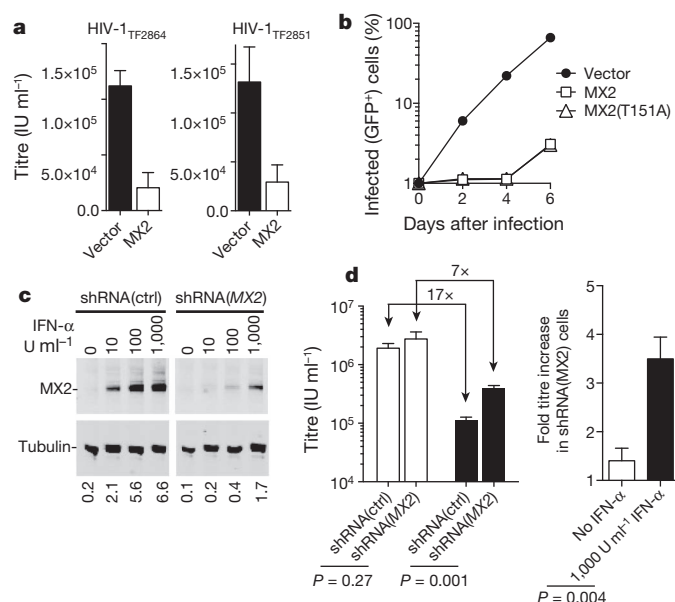


Figure 3 | MX2 inhibits replication-competent HIV-1 and is required for the full antiviral activity of IFN- α . **a**, Infection of empty vector (CSIB) or MX2-expressing GHOST-R5 cells with full-length primary HIV-1 strains. Titres are mean \pm s.d., $n = 3$ technical replicates, representative of two experiments. **b**, Growth of replication-competent HIV-1_{NL4-3} in empty vector (CSIB) or MX2-expressing GHOST-X4 cells (containing an HIV-2-LTR-GFP gene). **c**, Western blot analysis of MX2 and tubulin expression in IFN- α -treated THP-1 cells expressing control or MX2-targeted shRNAs. Numbers below each lane indicate fluorescence intensity associated with the MX2 band. **d**, HIV-1 GFP reporter virus infection of shRNA-expressing THP-1 cells from **c**, with (black) or without (white) IFN- α treatment. Titres are mean \pm s.d., $n = 3$ technical replicates, P values calculated using unpaired t -test, representative of three experiments.

to MX2 (Fig. 4b). Another mutation, G89V, which abolishes cyclophilin A binding by HIV-1 CA and the requirement for NUP358 during HIV-1 infection¹⁷, also conferred apparently complete MX2 resistance (Fig. 4b). Another CA mutation, N74D, which abolishes CA interaction with cleavage and polyadenylation specificity factor 6 (CPSF6)¹⁶, reduced but did not eliminate sensitivity to MX2, whereas the mutations G94D and A92E, which confer cyclophilin A sensitivity (cyclosporin A dependence) during early replication steps²⁰, slightly reduced MX2 sensitivity (Fig. 4b). These data demonstrate that the viral capsid governs the sensitivity of HIV-1 to MX2. In addition, they show that the antiviral activity of MX2 is specific, and unlikely to be the result of some generalized perturbation of cell physiology. Notably, the MX2-resistant CA mutant N57S exhibited a modest degree of resistance to IFN- α , relative to wild-type HIV-1, in THP-1 cells and HOS cells (Fig. 4c and Extended Data Fig. 7), supporting the notion that MX2 is one, but not the only, effector of the antiviral activity of IFN- α during the early steps of the HIV-1 replication cycle.

Because the cell-cycle-dependent HIV-1 CA mutant N57S was not inhibited by MX2 (Fig. 4b), we reasoned that arresting the cell cycle and thereby restricting HIV-1 infection to non-mitotic cells might potentiate the antiviral activity of MX2. Growth arrest of HOS or K562 cells with aphidicolin blocked infection by a control cell-cycle-dependent retrovirus (murine leukaemia virus) irrespective of MX2 expression, whereas HIV-1 was almost unaffected, as expected (Fig. 4d and Extended Data Fig. 8a, b). However, the inhibitory activity of MX2 was increased in non-dividing cells (Fig. 4d and Extended Data Fig. 8), in which it inhibited a single cycle of replication by ~ 30 -fold. In other words, MX2 both inhibited and conferred a degree of cell cycle dependence on wild-type HIV-1 infection.

Type I IFN inhibits HIV-1 replication at multiple points in the life cycle, both before and after the point at which MX2 seems to act^{2,6,12}.

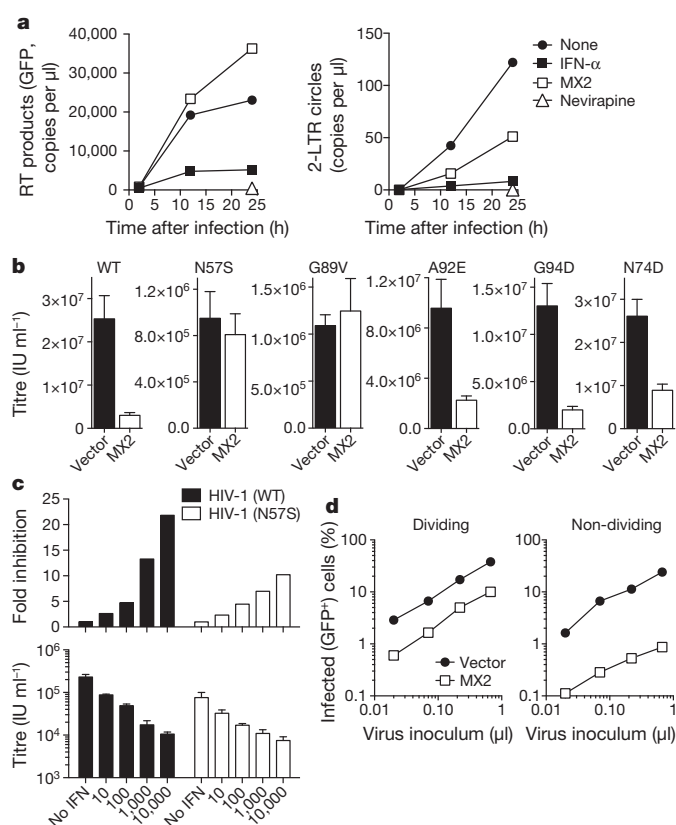


Figure 4 | MX2 activity reduces levels of nuclear HIV-1 DNA, is capsid dependent and is more potent in non-dividing cells. **a**, Quantitative PCR analysis of reverse transcript (RT, left) and 2-LTR circle (right) abundance in inhibitor-treated or MX2-expressing HOS cells. **b**, Wild-type (WT) or CA-mutant HIV-1-GFP reporter virus infection of vector or MX2-expressing HOS cells. Titres are mean \pm s.d., $n = 3$ technical replicates, representative of four experiments. **c**, Infectivity of wild-type and N57S CA-mutant HIV-1-GFP reporter viruses in untreated and IFN- α -treated THP-1 cells. Titres are mean \pm s.d., $n = 3$ technical replicates, representative of four experiments. Fold inhibition is the ratio of the mean titres on untreated and IFN- α -treated cells. **d**, HIV-1-GFP reporter virus infection of dividing and non-dividing (aphidicolin-treated) vector- or MX2-expressing HOS cell clones.

Thus MX2 is one of multiple effectors that contribute to the overall anti-HIV-1 activity of type I IFN. A few potential mechanisms might underlie the anti-HIV-1 activity of MX2. First, MX2 might directly target the incoming viral capsid, in a manner akin to the primate TRIM5- α and murine Fv1 antiretroviral proteins^{2,9,21}, or mutant cytoplasmic forms of CPSF6 (ref. 16). As with MX2, one consequence of the action of these capsid-targeting proteins is inhibition of the import of viral DNA into the nucleus, and in some cases their potency is enhanced in non-dividing cells^{16,22}. A second possibility is that MX2 inhibits particular nuclear import pathways, without regard to the precise nature of the import cargo, as mutant forms of MX2 have been shown to inhibit the nuclear accumulation of model cargos unrelated to HIV-1 (ref. 11). A third possibility is that MX2 acts after nuclear entry to destabilize viral DNA and/or inhibit integration. In these scenarios, CA mutations (G89V, N57S) could confer resistance by inhibiting interaction with MX2, by modulating the timing or extent of capsid uncoating, or by directing HIV-1 to alternative nuclear entry pathways. We note that the MX2-resistant G89V and N57S mutants exhibit reduced infectiousness in human cells, raising the possibility that the mutations abolish the use of pathways or processes during infection that are inhibited by MX2. Finally, it is possible that MX2 acts indirectly, for example by affecting the nuclear-cytoplasmic distribution of other cellular proteins that can interact with the viral capsid. However, the poor correlation in the degree of MX2 (Fig. 4) and CPSF6

(ref. 23) resistance/sensitivity exhibited by HIV-1 CA mutants suggests that redistribution of CPSF6 is unlikely to underlie the antiviral action of MX2. Although further work will be required to precisely define the molecular mechanisms involved, our findings demonstrate that MX2 is an effector in the anti-HIV-1 activity of type I IFN and underscore the remarkable diversity of proteins that cells can mobilize as antiretroviral defences.

METHODS SUMMARY

Gene expression in monocytoid cell lines was measured using human HT12 Expression Beadchip (Illumina) containing ~48,000 transcript probes, according to the manufacturer's instructions. Candidate antiviral genes, *MX2* and *MX2* mutants were expressed in K562, HOS or GHOST cells using the HIV-1-based vectors SCRPSY (which encodes TagRFP and puromycin resistance) or CSIB (which confers blasticidin resistance). *MX2*- and control-vector-expressing cells were used as populations or as single-cell clones in infection assays to evaluate *MX2* antiviral activity.

All single-cycle GFP reporter viruses were pseudotyped with VSV-G. Virus stocks were generated by transfecting 293T cells with Env-defective proviral DNA that encoded GFP in place of the *nef* gene, or in the case of primary HIV-1 strains, full-length proviral plasmids. Alternatively, packageable GFP-expressing retroviral vector and Gag-Pol packaging plasmids were cotransfected. Target cells in microwell plates were challenged with various doses of virus and single-cycle replication evaluated after 2 days. The proportion of cells infected with GFP reporter viruses, or replication-competent virus infection in GHOST cells (which contain an LTR-GFP indicator gene) in single cycle or spreading replication assays was measured by flow cytometry. *MX2* expression was reduced in target cells using a modified lentiviral shRNA expression vector (Origene). Non-dividing target cells were generated by aphidicolin treatment for 24 h before and during infection.

The abundance of viral DNA species was measured using quantitative PCR with primers directed to the GFP reporter gene, or to viral LTR sequences that are proximate only in 2-LTR circles. Western blotting was done using fluorescent antibodies and signals quantitated with a LI-COR Odyssey scanner. Deconvolution microscopy and image analysis was done using a Deltavision microscopy suite.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 August; accepted 12 September 2013.

Published online 13 October 2013.

- Ho, D. D. *et al.* Recombinant human interferon alpha-A suppresses HTLV-III replication *in vitro*. *Lancet* **325**, 602–604 (1985).
- Neil, S. & Bieniasz, P. Human immunodeficiency virus, restriction factors, and interferon. *J. Interferon Cytokine Res.* **29**, 569–580 (2009).
- Bitzegeio, J., Sampias, M., Bieniasz, P. D. & Hatzioannou, T. Adaptation to the interferon-induced antiviral state by human and simian immunodeficiency viruses. *J. Virol.* **87**, 3549–3560 (2013).
- Goujon, C. *et al.* Evidence for IFN α -induced, SAMHD1-independent inhibitors of early HIV-1 infection. *Retrovirology* **10**, 23 (2013).
- Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646–650 (2002).
- Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).
- Haller, O., Staeheli, P. & Kochs, G. Interferon-induced Mx proteins in antiviral host defense. *Biochimie* **89**, 812–818 (2007).
- Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485 (2011).
- Stremmlau, M. *et al.* The cytoplasmic body component TRIM5 α restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848–853 (2004).
- Melén, K. *et al.* Human MxB protein, an interferon- α -inducible GTPase, contains a nuclear targeting signal and is localized in the heterochromatin region beneath the nuclear envelope. *J. Biol. Chem.* **271**, 23478–23486 (1996).
- King, M. C., Raposo, G. & Lemmon, M. A. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. *Proc. Natl Acad. Sci. USA* **101**, 8957–8962 (2004).
- Goujon, C. & Malim, M. H. Characterization of the alpha interferon-induced postentry block to HIV-1 infection in primary human macrophages and T cells. *J. Virol.* **84**, 9254–9266 (2010).
- Yamashita, M., Perez, O., Hope, T. J. & Emerman, M. Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. *PLoS Pathog.* **3**, e156 (2007).
- Yamashita, M. & Emerman, M. Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. *J. Virol.* **78**, 5670–5678 (2004).
- Dismuke, D. J. & Aiken, C. Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. *J. Virol.* **80**, 3712–3720 (2006).
- Lee, K. *et al.* Flexible use of nuclear import pathways by HIV-1. *Cell Host Microbe* **7**, 221–233 (2010).
- Schaller, T. *et al.* HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. *PLoS Pathog.* **7**, e1002439 (2011).
- Koh, Y. *et al.* Differential effects of human immunodeficiency virus type 1 capsid and cellular factors nucleoporin 153 and LEDGF/p75 on the efficiency and specificity of viral DNA integration. *J. Virol.* **87**, 648–658 (2013).
- Rihn, S. J. *et al.* Extreme genetic fragility of the HIV-1 capsid. *PLoS Pathog.* **9**, e1003461 (2013).
- Sokolskaja, E., Sayah, D. M. & Luban, J. Target cell cyclophilin A modulates human immunodeficiency virus type 1 infectivity. *J. Virol.* **78**, 12800–12808 (2004).
- Stoye, J. P. *Fv1*, the mouse retrovirus resistance gene. *Rev. Sci. Tech.* **17**, 269–277 (1998).
- Yamashita, M. & Emerman, M. Cellular restriction targeting viral capsids perturbs human immunodeficiency virus type 1 infection of nondividing cells. *J. Virol.* **83**, 9835–9843 (2009).
- De Iaco, A. *et al.* TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. *Retrovirology* **10**, 20 (2013).

Acknowledgements We thank members of The Rockefeller University Genomics Resource Center for assistance with the microarray experiments and members of the Bieniasz laboratory for discussion and advice. This work was supported by grants from the National Institutes of Health; R37AI64003 (to P.D.B.), R01AI078788 (to T.H.) R01AI100720 (to M.Y.), AI091707 to C.M.R., AI057158 (to I. Lipkin, Northeast Biodefense Center, subcontracted to C.M.R.) and DK095031 to J.W.S., the Greenberg Medical Research Institute and the Starr Foundation (C.M.R.) and by the Howard Hughes Medical Institute.

Author Contributions M.K., S.S.Y., J.B., S.B.K., T.Z. and S.J.W. designed and executed the experiments and analysed the data. J.W.S. and C.M.R. provided an interferon-stimulated gene library and advice. M.Y. provided reagents and advice. T.H. provided reagents and advice and supervised the work. P.D.B. conceived the study, supervised the work and wrote the paper, with additional input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.D.B. (pbienias@adarc.org).

α TAT1 catalyses microtubule acetylation at clathrin-coated pits

Guillaume Montagnac^{1,2}, Vannary Meas-Yedid³, Marie Irondelle^{1,2}, Antonio Castro-Castro^{1,2}, Michel Franco⁴, Toshinobu Shida⁵, Maxence V. Nachury⁵, Alexandre Benmerah^{6,7}, Jean-Christophe Olivo-Marin³ & Philippe Chavrier^{1,2}

In most eukaryotic cells microtubules undergo post-translational modifications such as acetylation of α -tubulin on lysine 40, a widespread modification restricted to a subset of microtubules that turns over slowly¹. This subset of stable microtubules accumulates in cell protrusions² and regulates cell polarization³, migration and invasion^{4–7}. However, mechanisms restricting acetylation to these microtubules are unknown. Here we report that clathrin-coated pits (CCPs) control microtubule acetylation through a direct interaction of the α -tubulin acetyltransferase α TAT1 (refs 8, 9) with the clathrin adaptor AP2. We observe that about one-third of growing microtubule ends contact and pause at CCPs and that loss of CCPs decreases lysine 40 acetylation levels. We show that α TAT1 localizes to CCPs through a direct interaction with AP2 that is required for microtubule acetylation. In migrating cells, the polarized orientation of acetylated microtubules correlates with CCP accumulation at the leading edge¹⁰, and interaction of α TAT1 with AP2 is required for directional migration. We conclude that microtubules contacting CCPs become acetylated by α TAT1. In migrating cells, this mechanism ensures the acetylation of microtubules oriented towards the leading edge, thus promoting directional cell locomotion and chemotaxis.

Clathrin-mediated endocytosis is a fundamental process that regulates a wide variety of cell functions including signalling, migration and cell division. In migrating cells CCPs are asymmetrically distributed¹⁰

and endocytic carriers are enriched at the leading edge, probably providing a mechanism for rapid turnover of membrane components required for lamellipodia and adhesion site dynamics^{11,12}. In addition, close contacts between CCPs and microtubules have been reported¹³, although the functional consequences of these interactions have remained elusive. Here we set out to investigate the interaction between CCPs and the stable subset of microtubules that are oriented in the direction of protrusion.

Using total internal reflection fluorescence microscopy (TIRFM) we observed that a large proportion of green fluorescent protein (GFP)-end binding protein 1 (EB1)-labelled growing microtubule (+) ends disappeared upon contact with CCP labelled with monomeric red fluorescent protein-tagged clathrin light chain (mRFP-LCa) (Fig. 1a). Automated tracking and statistical co-localization analysis revealed that 31% of disappearances occurred when an EB1-positive comet contacted a CCP in HeLa cells (Fig. 1b), whereas the remaining comets disappeared in CCP-free regions. This percentage was significantly higher than prediction given by random superposition of disappearing EB1 events and CCPs (Fig. 1b and Methods). Approximately 28% of growing GFP- α -tubulin-labelled microtubule ends that passed over a CCP paused at this structure in MDA-MB-231 cells (Fig. 1c), similar to the 27% of EB1 comets that stopped at CCPs in these cells (Fig. 1b); the pause time was highly variable with an average of 16.8 ± 15.1 s (mean \pm s.e.m.). When

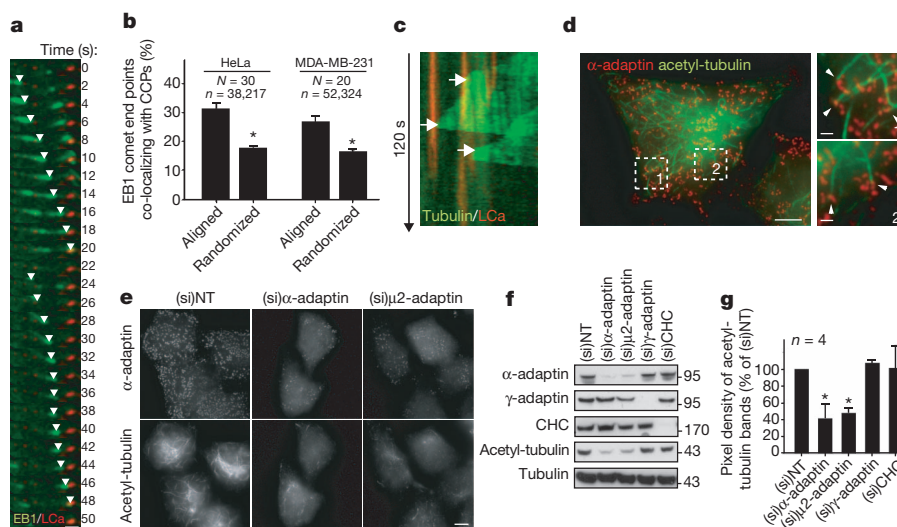


Figure 1 | Microtubules pause at CCPs and are acetylated in an AP2-dependent manner. **a**, **b**, GFP-EB1 comets stopping at CCPs (**a**, TIRFM, HeLa cells) and quantification (**b**, see Methods; *N*, number of cells; *n*, number of EB1 comets). **c**, GFP-tubulin-positive microtubule contacting CCP. **d**, **e**, Control (**d**) or siRNA (si)-treated (**e**) HeLa cells stained for α -adaptin and

K40-acetyl-tubulin. **f**, **g**, Protein expression in HeLa cells treated with the indicated siRNAs (molecular weights in kDa). Quantification in percentage \pm s.e.m. of non-targeting siRNAs ((si)NT), **P* < 0.001. Scale bars, 10 μ m, and 2 μ m in insets.

¹Institut Curie, Research Center, 75005 Paris, France. ²Membrane and Cytoskeleton Dynamics, CNRS UMR 144, 75005 Paris, France. ³Unité d'Analyse d'Images Quantitative, Institut Pasteur, CNRS URA 2582, 75015 Paris, France. ⁴Institut de Pharmacologie Moléculaire et Cellulaire, CNRS UMR 6097, Université de Nice-Sophia Antipolis, 06560 Valbonne, France. ⁵Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, California 94305-5345, USA. ⁶INSERM, U983, Hôpital Necker-Enfants Malades, 75015 Paris, France. ⁷Université Paris Descartes, Sorbonne Paris Cité, Institut Imagine, 75015 Paris, France.

CCPs were disrupted by silencing the α -adaptin subunit of AP2, EB1 comets travelled significantly longer distances ($\sim 2.6 \mu\text{m}$ compared to $\sim 2 \mu\text{m}$ in control cells; Extended Data Fig. 1). Collectively, these data indicate that microtubules can pause and anchor transiently at CCPs.

Anchoring events at the cell periphery have a role in the generation of stable microtubules¹⁴. We found that $17 \pm 0.7\%$ (mean \pm s.e.m.) of the visible acetylated microtubule extremities were in contact with CCPs at the cell periphery (Fig. 1d). Notably, the amount of lysine 40 (K40)-acetylated tubulin was markedly reduced when CCPs were disrupted, although there was no global change in microtubule network organization or on levels of other microtubule modifications (Fig. 1e–g and Extended Data Figs 2a and 3). In addition, depletion of AP1 subunit γ -adaptin did not affect K40 acetylation (Fig. 1f, g and Extended Data Fig. 2b). Silencing of clathrin heavy chain (CHC), which did not modify AP2 localization at the plasma membrane (ref. 15 and Extended Data Fig. 2), did not affect K40 acetylation levels (Fig. 1f, g and Extended Data Fig. 2b), suggesting that the role of AP2 in microtubule acetylation is independent of endocytosis. Thus, we conclude that there is a positive correlation between CCP density and K40 acetylation levels. In addition, loss of dynamin function was reported to increase the density of CCPs at the plasma membrane and to enhance tubulin acetylation^{16,17}.

Using TIRFM, we observed that GFP- α TAT1 accumulated into CCPs in a microtubule-independent manner and that $\sim 35\%$ of CCPs (302 of 853 CCPs analysed from 7 different cells) were positive for endogenous α TAT1 (Fig. 2a and Extended Data Fig. 4b, c). GFP- α TAT1 was also found in focal adhesions and was associated with the microtubule network (Extended Data Fig. 4a, d). In addition, co-immunoprecipitation assay with GFP- α TAT1 recovered tubulin as well as α -adaptin but not

the AP1 subunit γ -adaptin (Fig. 2b). Notably, α TAT1 knockdown did not affect microtubule pausing at CCPs nor clathrin-mediated endocytosis (Extended Data Fig. 5a, b).

α TAT1 comprises a catalytic domain (residues 1–193) and an unstructured tail (residues 193–421; Extended Data Fig. 6a)⁹. We found that residues 307–387 contained the minimal binding sites for both AP2 and tubulin (Fig. 2c). This region interacted directly with purified AP2 and tubulin (Fig. 2d); there was no competition between tubulin and AP2 for binding to glutathione S-transferase (GST)- α TAT1(307–387), indicating that the two proteins have distinct binding sites (Fig. 2d). Interestingly, GST- α TAT1(307–387) did not interact with the recombinant AP2 ‘core complex’ lacking the hinge and appendage domains of α - and β 2-adaptin¹⁸ (Extended Data Fig. 6c). In addition, the AP2-binding region of α TAT1 (307–387) pulled down full-length α -adaptin and a truncated variant lacking the appendage domain (residues 1–690), but not a construct missing both the hinge and appendage domains (residues 1–620; Fig. 2e and Extended Data Fig. 6b). Conversely, the hinge and appendage domains (residues 603–938) were robustly pulled down by GST- α TAT1(307–387) (Fig. 2e). Together, these data support the conclusion that α TAT1 associates directly with the hinge domain of α -adaptin. Notably, a shorter α TAT1 isoform lacking the AP2-binding domain (Extended Data Fig. 7a)³ associated neither with microtubules nor with CCPs (Extended Data Fig. 7b, c), suggesting that different α TAT1 isoforms are differentially localized. Consistent with an essential role for the interaction between α TAT1 and AP2 in K40 acetylation, expression of wild-type α -adaptin but not α -adaptin(1–620) restored K40 acetylation in α -adaptin-depleted HeLa cells (Fig. 2f, g).

We next investigated whether CCPs are sites of microtubule acetylation. When microtubules were depolymerized with nocodazole, K40-acetylated tubulin was barely detected (Extended Data Fig. 8a, b; time 0); only bright dots corresponding to centrosomes remained visible (Fig. 3a). In agreement with previous findings¹⁹, short acetylated microtubule segments were visible in the vicinity of the adherent plasma membrane 5 min after nocodazole-washout-induced microtubule regrowth, whereas acetylated-K40 levels increased (Fig. 3a and Extended Data Fig. 8a–d). Many of these acetylated segments were at the extremity of longer microtubules (Fig. 3b) and $\sim 24\%$ of these segments were either overlapping

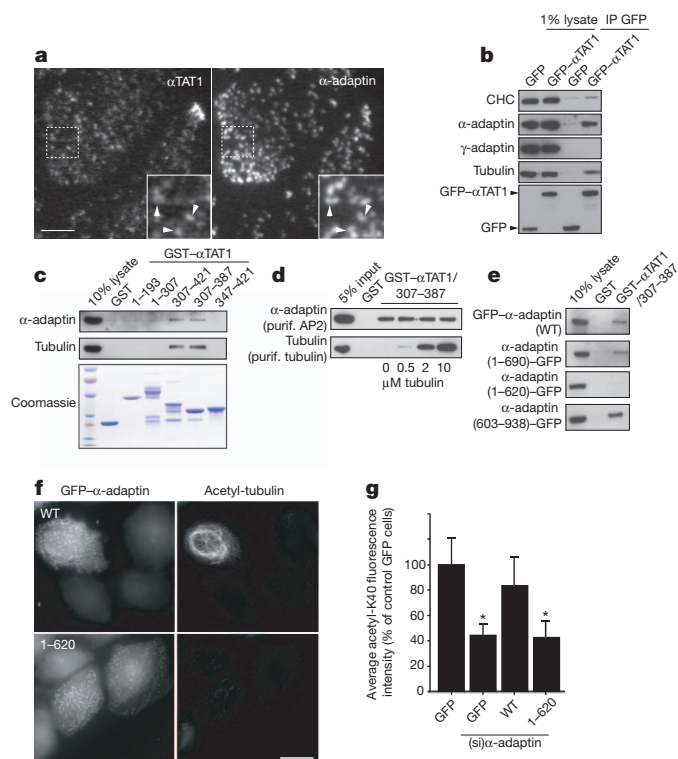


Figure 2 | α TAT1 interaction with AP2 is required for α -tubulin acetylation. **a**, HeLa cells stained for α TAT1 and α -adaptin (TIRFM). **b**, **c**, Immunoprecipitation (IP, **b**) or GST-pulldown (**c**) experiments of GFP- α TAT1 or GST- α TAT1 fragments, respectively, with HeLa cell lysate. **d**, *In vitro* direct binding assay between GST- α TAT1(307–387) and purified AP2 and tubulin. **e**, Pulldown assays of GST- α TAT1(307–387) with GFP-tagged α -adaptin variants from HeLa cell lysates. **f**, **g**, Acetylated-K40 levels in α -adaptin-depleted HeLa cells transfected with the indicated construct. Fluorescence intensity of acetylated-K40 expressed as percentage \pm s.e.m. of non-targeting siRNA-treated, GFP-transfected cells (* $P < 0.001$). Scale bars, 10 μm .

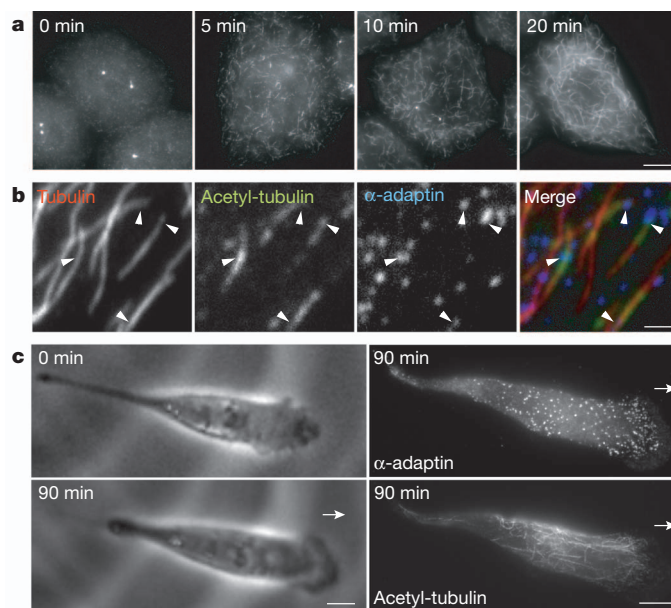


Figure 3 | Spatial restriction of microtubule acetylation by CCP distribution. **a**, **b**, HeLa cells stained for K40 acetyl-tubulin (**a**) or α -adaptin, total tubulin and acetylated-K40 (**b**) at the indicated times (**a**) or 5 min (**b**) after nocodazole washout. **c**, Live MDA-MB-231 cells imaged for 90 min (left) and then fixed and stained for α -adaptin and K40 acetyl-tubulin (right). Scale bars, 10 μm (**a** and **c**) and 2 μm (**b**).

with or had one end in contact with a CCP (Fig. 3b, as compared to 15% when CCP distribution was randomized; $P < 0.001$, Chi-squared test). These data are consistent with the conclusion that CCPs are sites of microtubule acetylation, although we do not exclude that other microtubule-acetylation sites may exist. The length of acetylated microtubule segments increased progressively with time, concomitant with acetylated K40 levels also reaching a plateau, whereas K40 acetylation was strongly delayed and reduced in α -adaplin-depleted cells (Fig. 3a and Extended Data Fig. 8a, b). Because Golgi-associated microtubules are rapidly acetylated under nocodazole-washout conditions²⁰, the subset of microtubule being acetylated at CCPs 5 min after nocodazole washout could arise from Golgi-mediated nucleation²¹.

Stable microtubules oriented towards the leading edge have a role in migrating cells^{22,23}. We observed that in MDA-MB-231 cells migrating on a two-dimensional substrate or through a three-dimensional matrix of type I collagen fibres, acetylated microtubules were oriented towards the cell front where CCPs accumulated¹⁰ (Fig. 3c and Extended Data Fig. 9a, b). Moreover, acetylated-K40 levels were reduced in AP2- or α TAT1-depleted cells within the three-dimensional environment (Extended Data Fig. 9c). This suggested that α TAT1 associated with CCPs at

the cell front controls the polarized distribution of acetylated microtubules. On a two-dimensional substrate, MDA-MB-231 cells depleted for α -adaplin or α TAT1 moved with a similar velocity to control cells but in less linear paths, indicating that both proteins regulate the directionality of cell migration (Extended Data Fig. 10a, b). By contrast, CHC depletion reduced velocity but did not affect directionality (Extended Data Fig. 10a, b), possibly reflecting the AP2-independent role of clathrin in focal adhesion turnover, a process that is required for migration²⁴. Inactivation of AP2 or α TAT1 also inhibited the invasive migration of MDA-MB-231 cells in a three-dimensional environment as potently as knockdown of the pro-invasive metalloproteinase MT1-MMP (Fig. 4a and Extended Data Fig. 10c)^{6,25}. Migration of cancer cells away from the primary tumour is generally oriented towards growth factors in the microenvironment²⁶. We generated a gradient of epidermal growth factor (EGF) in the three-dimensional collagen gel (Fig. 4b–d) and observed that although the intensity of the gradient progressively diminished over time (Fig. 4e), the slope remained approximately constant (Fig. 4f) and cells effectively moved towards the gradient (Fig. 4g, h). Silencing of AP2 or α TAT1, but not of CHC, inhibited cell movement towards the EGF gradient (Fig. 4g–h) as well as the directionality (persistence)

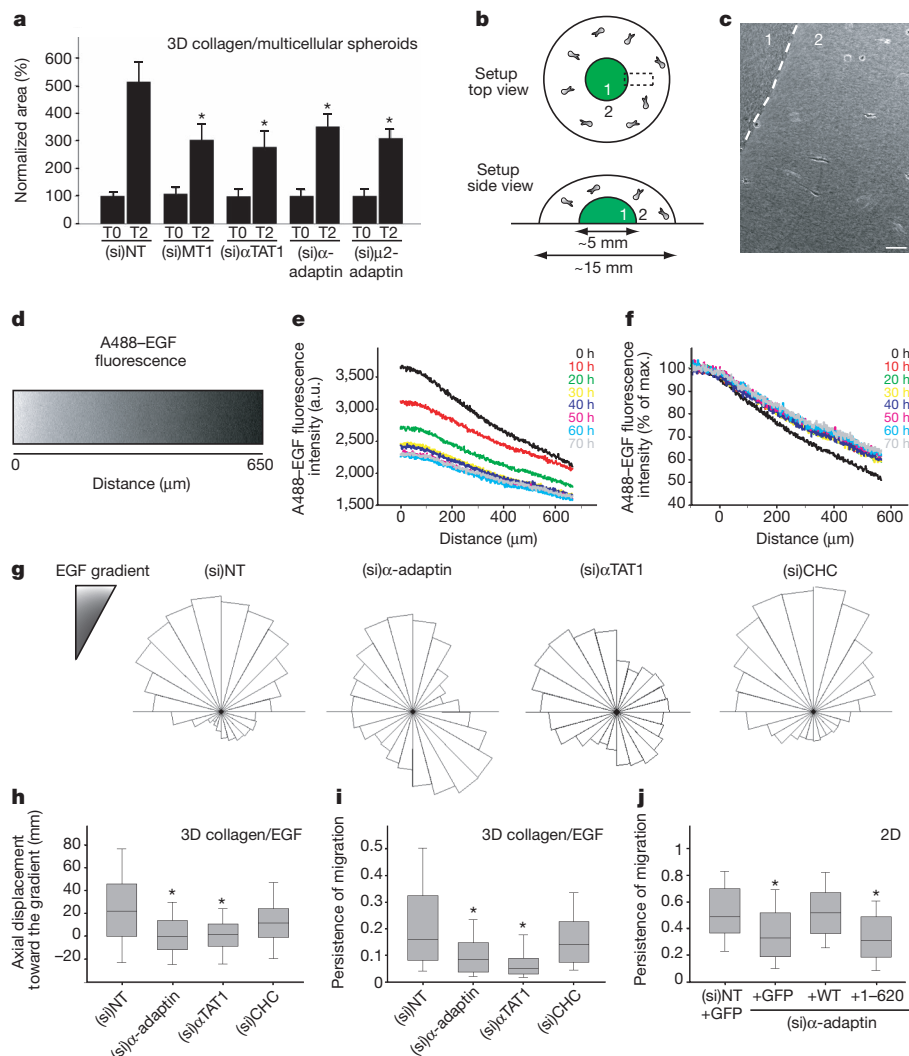


Figure 4 | α TAT1 interaction with AP2 is required for directed cell migration. **a**, Area of three-dimensional (3D) invasion after 2 days (T2) by spheroids of siRNAs-treated MDA-MB-231 cells. **b**, **c**, Schematic representation (**b**) or phase-contrast image (**c**) of the three-dimensional collagen I EGF-chemotaxis setup. Scale bar, 50 μ m. **d**, Alexa 488 (A488)–EGF gradient at time 0 in a region corresponding to boxed area in **b**. **e–f**, Evolution over time of the intensity (**e**) and slope (**f**) of the A488-EGF gradient.

g, Angular distribution relative to gradient orientation of siRNA-treated MDA-MB-231 cells. **h**, **i**, Axial displacement towards the gradient (**h**) and persistence of migration (**i**) of siRNA-treated MDA-MB-231 cells. **j**, Persistence of migration on glass coverslip of MDA-MB-231 cells depleted for α -adaplin and expressing the indicated constructs. Error bars indicate mean \pm s.e.m. * $P < 0.001$ in **a** and $P < 0.05$ in **h–j**.

of migration (Fig. 4i). Finally, consistent with an essential role for the interaction between α TAT1 and AP2, expression of wild-type α -adaptin but not α -adaptin(1–620) restored the directionality of two-dimensional migration in α -adaptin-depleted cells (Fig. 4j).

In conclusion, we report an unanticipated role for CCPs in microtubule acetylation through a direct interaction between α TAT1 and AP2. We propose that the asymmetric distribution of CCPs shapes the acetylated microtubule network by selectively acetylating microtubules that are oriented towards the leading edge, thus promoting directed cell motility and chemotaxis.

METHODS SUMMARY

For live-cell TIRFM and spinning disk microscopy, cells were imaged for 100 ms at 1-s intervals for 120 s. Automatic detection and tracking of fluorescent EB1 and LCa spots was performed using the ICY software²⁷. For siRNA depletion, HeLa cells or MDA-MB-231 cells were transfected with indicated siRNAs by using Oligofectamine (Invitrogen) or Lullaby (OZ Biosciences), respectively. DNA constructs encoding GFP-tagged full-length α TAT1 (residues 1–421) or GST-tagged α TAT1 variants (residues 1–193, 1–307, 307–421, 307–387 or 347–421) were obtained by PCR by using murine α TAT1 complementary DNA as a template and subcloning into pEGFP-C3 (Clontech) or into pGEX4T1 (Amersham Pharmacia Biotech), respectively. GST constructs were expressed in BL21 *Escherichia coli* and purified by using glutathione-Sepharose beads (GE Healthcare). Pulldown experiments were performed by incubating GST- α TAT1 domains with HeLa cell lysate prepared in 50 mM Tris, pH 7.4, 137 mM NaCl, 1 mM MgCl₂, 10% glycerol, 1% Triton X-100 with protease inhibitors. Immunoprecipitation assays were performed by incubating lysates of HeLa cells expressing GFP or GFP- α TAT1 with GFP-Trap-coupled agarose beads (ChromoTek). To analyse microtubule regrowth and K40 acetylation, HeLa cells were first incubated with 10 μ M nocodazole for 5 h before washing out the drug with fresh medium. The three-dimensional collagen chemotaxis setup was built by polymerizing a 50- μ l collagen-I (acid extracted) droplet containing EGF followed by polymerization of a 200- μ l EGF-free collagen-I droplet containing 10,000 cells per ml atop the inner gel. Migration of MDA-MB-231 cells was analysed by manual tracking using MetaMorph software.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 July 2012; accepted 9 August 2013.

Published online 6 October 2013.

- Perdiz, D., Mackeh, R., Pous, C. & Baillet, A. The ins and outs of tubulin acetylation: more than just a post-translational modification? *Cell. Signal.* **23**, 763–771 (2011).
- Wloga, D. & Gaertig, J. Post-translational modifications of microtubules. *J. Cell Sci.* **123**, 3447–3455 (2010).
- Witte, H., Neukirchen, D. & Bradke, F. Microtubule stabilization specifies initial neuronal polarization. *J. Cell Biol.* **180**, 619–632 (2008).
- Castro-Castro, A., Janke, C., Montagnac, G., Paul-Gilloteaux, P. & Chavrier, P. ATAT1/MEC-17 acetyltransferase and HDAC6 deacetylase control a balance of acetylation of alpha-tubulin and cortactin and regulate MT1-MMP trafficking and breast tumor cell invasion. *Eur. J. Cell Biol.* **91**, 950–960 (2012).
- Hubbert, C. *et al.* HDAC6 is a microtubule-associated deacetylase. *Nature* **417**, 455–458 (2002).
- Rey, M., Irondele, M., Waharte, F., Lizarraga, F. & Chavrier, P. HDAC6 is required for invadopodia activity and invasion by breast tumor cells. *Eur. J. Cell Biol.* **90**, 128–135 (2011).
- Tran, A. D. *et al.* HDAC6 deacetylation of tubulin modulates dynamics of cellular adhesions. *J. Cell Sci.* **120**, 1469–1479 (2007).
- Akella, J. S. *et al.* MEC-17 is an α -tubulin acetyltransferase. *Nature* **467**, 218–222 (2010).
- Shida, T., Cueva, J. G., Xu, Z., Goodman, M. B. & Nachury, M. V. The major α -tubulin K40 acetyltransferase α TAT1 promotes rapid ciliogenesis and efficient mechanosensation. *Proc. Natl Acad. Sci. USA* **107**, 21517–21522 (2010).
- Rappoport, J. Z. & Simon, S. M. Real-time analysis of clathrin-mediated endocytosis during cell migration. *J. Cell Sci.* **116**, 847–855 (2003).
- Caswell, P. T. *et al.* Rab25 associates with α 5B1 integrin to promote invasive migration in 3D microenvironments. *Dev. Cell* **13**, 496–510 (2007).
- Howes, M. T. *et al.* Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells. *J. Cell Biol.* **190**, 675–691 (2010).
- Rappoport, J. Z., Taha, B. W. & Simon, S. M. Movement of plasma-membrane-associated clathrin spots along the microtubule cytoskeleton. *Traffic* **4**, 460–467 (2003).
- Gundersen, G. G. Microtubule capture: IQGAP and CLIP-170 expand the repertoire. *Curr. Biol.* **12**, 645–647 (2002).
- Hinrichsen, L., Harborth, J., Andrees, L., Weber, K. & Ungewickell, E. J. Effect of clathrin heavy chain- and α -adaptin-specific small inhibitory RNAs on endocytic accessory proteins and receptor trafficking in HeLa cells. *J. Biol. Chem.* **278**, 45160–45170 (2003).
- Ferguson, S. M. *et al.* Coordinated actions of actin and BAR proteins upstream of dynamin at endocytic clathrin-coated pits. *Dev. Cell* **17**, 811–822 (2009).
- Tanabe, K. & Takei, K. Dynamic instability of microtubules requires dynamin 2 and is impaired in a Charcot-Marie-Tooth mutant. *J. Cell Biol.* **185**, 939–948 (2009).
- Collins, B. M., McCoy, A. J., Kent, H. M., Evans, P. R. & Owen, D. J. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109**, 523–535 (2002).
- Bulinski, J. C., Richards, J. E. & Piperno, G. Posttranslational modifications of alpha tubulin: dephosphorylation and acetylation differentiate populations of interphase microtubules in cultured cells. *J. Cell Biol.* **106**, 1213–1220 (1988).
- Chabin-Brion, K. *et al.* The Golgi complex is a microtubule-organizing organelle. *Mol. Biol. Cell* **12**, 2047–2060 (2001).
- Efimov, A. *et al.* Asymmetric CLASP-dependent nucleation of noncentrosomal microtubules at the trans-Golgi network. *Dev. Cell* **12**, 917–930 (2007).
- Gundersen, G. G. & Bulinski, J. C. Selective stabilization of microtubules oriented toward the direction of cell migration. *Proc. Natl Acad. Sci. USA* **85**, 5946–5950 (1988).
- Watanabe, T., Noritake, J. & Kaibuchi, K. Regulation of microtubules in cell migration. *Trends Cell Biol.* **15**, 76–83 (2005).
- Ezratty, E. J., Bertaux, C., Marcantonio, E. E. & Gundersen, G. G. Clathrin mediates integrin endocytosis for focal adhesion disassembly in migrating cells. *J. Cell Biol.* **187**, 733–747 (2009).
- Rowe, R. G. & Weiss, S. J. Breaching the basement membrane: who, when and how? *Trends Cell Biol.* **18**, 560–574 (2008).
- Condeelis, J. & Pollard, J. W. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell* **124**, 263–266 (2006).
- de Chaumont, F. *et al.* Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods* **9**, 690–696 (2012).

Acknowledgements The authors wish to thank P. Tran and C. Janke for comments on the manuscript and S. Linder for the suggestion of the three-dimensional collagen I EGF-chemotaxis assay. We thank E. Macia for purification of recombinant AP2 complex and S. Lemeer for generation of α -adaptin mutants. We thank the Cell and Tissue Imaging Facility and Nikon Imaging Center@Institut Curie & Centre National de la Recherche Scientifique (CNRS) for help with image acquisition. Core funding for this work was provided by the Institut Curie and the CNRS and additional support was provided by grants from Fondation ARC pour la Recherche contre le Cancer (SL220100601356) and Institut National du Cancer (2009-1-PL BIO-12-IC-1) to P.C.

Author Contributions G.M. designed the project and the experiments, performed experiments, analysed results and wrote the manuscript. V.M.-Y. and J.-C.O.-M. generated software for automated tracking analyses. M.I. and A.C.-C. performed and quantified multicellular spheroid three-dimensional migration experiments. M.F. purified proteins and designed experiments. T.S., M.V.N. and A.B. provided critical materials and designed experiments. P.C. supervised the study, contributed to experimental design and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.M. (guillaume.montagnac@curie.fr) or P.C. (philippe.chavrier@curie.fr).

Microbial production of short-chain alkanes

Yong Jun Choi¹ & Sang Yup Lee^{1,2}

Increasing concerns about limited fossil fuels and global environmental problems have focused attention on the need to develop sustainable biofuels from renewable resources. Although microbial production of diesel has been reported, production of another much in demand transport fuel, petrol (gasoline), has not yet been demonstrated. Here we report the development of platform *Escherichia coli* strains that are capable of producing short-chain alkanes (SCAs; petrol), free fatty acids (FFAs), fatty esters and fatty alcohols through the fatty acyl (acyl carrier protein (ACP)) to fatty acid to fatty acyl-CoA pathway. First, the β -oxidation pathway was blocked by deleting the *fadE* gene to prevent the degradation of fatty acyl-CoAs generated *in vivo*. To increase the formation of short-chain fatty acids suitable for subsequent conversion to SCAs *in vivo*, the activity of 3-oxoacyl-ACP synthase (FabH)¹, which is inhibited by unsaturated fatty acyl-ACPs², was enhanced to promote the initiation of fatty acid biosynthesis by deleting the *fadR* gene; deletion of the *fadR* gene prevents upregulation of the *fabA* and *fabB* genes responsible for unsaturated fatty acids biosynthesis³. A modified thioesterase⁴ was used to convert short-chain fatty acyl-ACPs to the corresponding FFAs, which were then converted to SCAs by the sequential reactions of *E. coli* fatty acyl-CoA synthetase, *Clostridium acetobutylicum* fatty acyl-CoA reductase and *Arabidopsis thaliana* fatty aldehyde decarboxylase. The final engineered strain produced up to 580.8 mg l⁻¹ of SCAs consisting of nonane (327.8 mg l⁻¹), dodecane (136.5 mg l⁻¹), tridecane (64.8 mg l⁻¹), 2-methyl-dodecane (42.8 mg l⁻¹) and tetradecane (8.9 mg l⁻¹), together with small amounts of other hydrocarbons. Furthermore, this platform strain could produce short-chain FFAs using a *fadD*-deleted strain, and short-chain fatty esters by introducing the *Acinetobacter* sp. ADP1 wax ester synthase (*atfA*)⁵ and the *E. coli* mutant alcohol dehydrogenase (*adhE^{mut}*)⁶.

Bio-based sustainable production of fuels has been attracting increasing interest for our sustainable future⁷. Hydrocarbon, such as alkane or alkene, is of particular interest owing to its potential to be used as an advanced biofuel that is similar to the petro-based fuels currently in use and superior to other biofuels in many aspects, including its high energy content (for example, it has a 30% higher energy content than ethanol)⁸. There have been a few reports on the bio-based production of C13–C17 long-chain hydrocarbons for substituting for diesel⁹. Microbial production of up to 300 mg l⁻¹ of long-chain hydrocarbons, mainly pentadecane and heptadecane, was achieved by using an engineered *E. coli* strain harbouring a cyanobacterial alkane biosynthesis operon encoding acyl-ACP reductase and aldehyde decarboxylase⁹. Another study also reported production of even or odd numbered long-chain alkanes in *E. coli* by the overexpression of the *Bacillus subtilis* *fabH* gene¹⁰. In these studies, hydrocarbons were produced by decarboxylation of fatty aldehydes, which are directly generated from fatty acyl-ACPs. More recently, long-chain alkanes were produced from fatty acids by using fatty acid reductase and aldehyde decarboxylase¹¹.

Petrol, a mixture of C4–C12 short-chain hydrocarbons (SCHCs)¹², is a liquid fuel primarily used in internal combustion engines. Although short-chain alcohols were produced to substitute for petrol^{13,14}, they are inferior to petrol in their fuel properties (Supplementary Table 1). Thus,

it is of great interest to produce SCHCs directly that have the potential to be used directly as petrol¹⁵. However, there has been no report so far about the production of such SCHCs by microbial fermentation. This seems to be because most of the bacterial fatty acids identified are C14–C18 long-chain ones. Here we report the development of engineered *E. coli* strains capable of producing SCAs suitable for petrol by engineering fatty acid biosynthesis and degradation pathways. This was achieved, in a different way from previous studies on the production of long-chain hydrocarbons, by introducing a new pathway involving a mutant fatty acyl-ACP thioesterase, fatty acyl-CoA synthetase, fatty acyl-CoA reductase and fatty aldehyde decarboxylase into engineered *E. coli* supporting generation of short-chain fatty acyl-ACPs. The detailed strategy for the production of SCAs is described in Fig. 1 and Supplementary Fig. 1. This strategy also allows production of short-chain FFAs, fatty esters and fatty alcohols as described below.

In the production of fatty-acid-based biofuels, FFAs derived from fatty acyl-ACPs by thioesterases are important intermediate metabolites. To examine the performance of different thioesterases, the *fadD* gene was deleted in *E. coli* strain W3110 to prevent conversion of FFAs to fatty acyl-CoAs. Among three thioesterases encoded by the *E. coli* *tesB* gene¹⁶, *E. coli* 'tesA (a leaderless version of *tesA*)' gene¹⁷, and the *Umbellularia californica* *fatB* gene¹⁸, the *fadD*-deleted W3110 strain expressing 'TesA was found to be the best, producing 313 mg l⁻¹ (Fig. 2a) of mixed FFAs (mainly C16 and small amounts of C8, C10, C12 and C14; Fig. 2b).

Because 'TesA preferentially hydrolyses long-chain fatty acyl-ACPs¹⁷, an engineered thioesterase capable of converting short-chain fatty acyl-ACPs to FFAs was needed. Because TesA with a L109P mutation showed hydrolytic activity on both short- and long-chain fatty acyl-ACPs⁴, 'TesA was similarly engineered to make 'TesA(L109P). Recombinant *E. coli* *fadD*-deleted W3110 expressing 'TesA(L109P) was able to produce a FFA mixture of short carbon lengths; C16 FFAs decreased by 91%, whereas C14, C12 and C10 FFAs increased by 6.8-fold, 12.8-fold and 2.2-fold, respectively (Fig. 2b). The percentages of C12 and C14 FFAs produced by *fadD*-deleted W3110 harbouring 'TesA(L109P) were 19.5% and 69.9%, respectively (Fig. 2b). Many thioesterases having different substrate specificities and activities¹⁹ can be similarly used. Other approaches recently reported can also be taken to produce short-chain FFAs^{20,21} (see Supplementary Discussion).

For the production of SCAs, the *fadE* gene needs to be deleted to block β -oxidation (Fig. 1a). Thus, the GAS1 strain was constructed by deleting the *fadE* gene in W3110. As in the *fadD*-deleted W3110 strain, the *fadD* gene was also deleted in the GAS1 strain to allow production of FFAs. The *fadD*-deleted GAS1 strain expressing 'TesA(L109P) was also able to produce short-chain FFAs (Fig. 2c); there were some variations in the composition of FFAs but C14 was the most prevalent one as in *fadD*-deleted W3110 expressing 'TesA(L109P).

Formation of short-chain FFAs can be enhanced by promoting the initiation of fatty acid biosynthesis, that is, the formation of β -ketoacyl-ACP by the condensation of acetyl-CoA and malonyl-ACP by 3-ketoacyl-ACP synthase (FabH)^{1,22}. The overexpression of the *fabH* gene indeed increased production of short-chain fatty acids; C14

¹Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 program), BioProcess Engineering Research Center and Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea.

²Bioinformatics and BioProcess Engineering Research Center, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, South Korea.

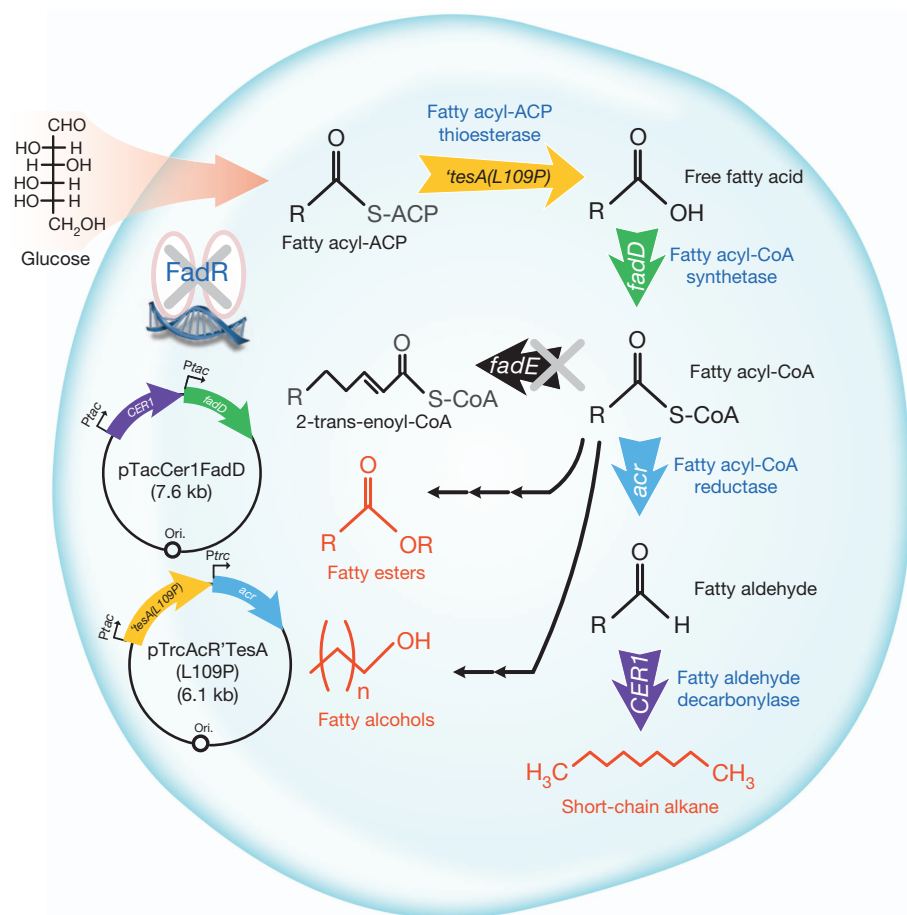


Figure 1 | Metabolic engineering of *E. coli* for the production of short-chain alkanes. The overall strategy for the production of short-chain alkanes (SCAs) is shown. The key enzymes are represented in the same colour. The genes knocked out are indicated by X. Glucose is converted into SCAs, fatty esters and fatty alcohols through the fatty acyl-ACP to fatty acid to fatty acyl-CoA pathway. The detailed metabolic engineering strategies used for the production of SCAs are described in Supplementary Fig. 1. The enzymes encoded by the genes shown are: *acr*, fatty acyl-CoA reductase; *CER1*, fatty aldehyde decarbonylase; *fadD*, fatty acyl-CoA synthetase; *fadE*, acyl-CoA dehydrogenase; *tesA*(L109P), leaderless mutant thioesterase I. FadR is a global transcriptional regulator protein.

FFAs decreased by 77.4%, whereas C8 and C10 FFAs increased by 65% and 16%, respectively, in the *fadD*-deleted W3110 strain expressing 'TesA(L109P) and FabH. Moreover, the *fadD*-deleted GAS1 strain expressing 'TesA(L109P) and FabH produced C8 FFA by 5.7-fold more than the *fadD*-deleted W3110 strain expressing 'TesA(L109P) (Supplementary Fig. 2). FabH is known to be inhibited by the unsaturated fatty acyl-ACPs². Because FadR positively regulates unsaturated fatty acid biosynthesis by upregulating the β -hydroxyacyl-ACP dehydratase and 3-ketoacyl-ACP synthase I operon (*fabAB*)³, the *fadR* gene was also deleted in GAS1 to make the GAS2 strain. The *fadD*-deleted GAS2 strain expressing 'TesA(L109P) produced 2.6-fold and 1.6-fold more C10 and C12 FFAs and 64.6% lower C14 FFAs compared with the *fadD*-deleted GAS1 strain expressing 'TesA(L109P) (Fig. 2c, d). Thus, either overexpression of the *fabH* gene or deletion of the *fadR* gene results in the enhanced production of short-chain FFAs. Combining the two approaches of *fadR* deletion and *fabH* amplification did not further improve the production of short-chain FFAs mainly because of growth retardation. The importance of FabH in increasing short-chain fatty acids was also evaluated by knockdown experiments using synthetic RNA (sRNA)²³. The knockdown of the *fabH* gene decreased the titre of short-chain fatty acids and increased the titre of long-chain fatty acids; C10 FFA decreased by 94%, whereas C14 FFAs increased by twofold in the *fadD*-deleted GAS2 strain (Supplementary Fig. 3). Thus, short-chain FFAs could be successfully produced in *E. coli* by deleting the *fadD*, *fadE* and *fadR* genes.

Next, conversion of short-chain FFAs to SCAs was performed in the GAS2 strain by amplifying the *fadD* gene and introducing the fatty acyl-CoA reductase and fatty aldehyde decarbonylase reactions (Fig. 1 and Supplementary Fig. 1). Even though *E. coli* is known to use only long-chain FFAs, recombinant *E. coli* overexpressing the *fadD* gene can utilize C8 and C10 FFAs²⁴. This suggests that FadD can transfer CoA to both long- and short-chain FFAs. Thus, the GAS3 strain was

constructed by replacing the native promoter of the *fadD* gene with the strong *trc* promoter in the chromosome of GAS2. Also, the FadD level was further increased by plasmid-based overexpression of *fadD*. Then the *Clostridium acetobutylicum* *acr* gene encoding a fatty acyl-CoA reductase for the reduction of fatty acyl-CoAs to fatty aldehydes and the *E. coli* codon-optimized *Arabidopsis thaliana* *CER1* gene encoding a fatty aldehyde decarbonylase²⁵ for the decarbonylation of fatty aldehydes to corresponding hydrocarbons were introduced by plasmid-based overexpression under the *trc* and *tac* promoters, respectively.

Fed-batch culture of the GAS3 strain harbouring pTacCer1FadD and pTrcAcrTesA(L109P) resulted in the production of 396.5 mg l⁻¹ of SCHCs, composed of 18.1 mg l⁻¹ octene, 34.1 mg l⁻¹ 2-octene, 217.0 mg l⁻¹ nonane, 100.0 mg l⁻¹ dodecane, 24.1 mg l⁻¹ tridecane and 3.2 mg l⁻¹ tetradecane (Fig. 3a, Supplementary Figs 4 and 5). To increase the titre of SCHCs further, the activity of fatty aldehyde decarbonylase (CER1) needed to be enhanced. The expression level of CER1 was found to be higher at 30 °C, compared with results obtained at other temperatures examined (Supplementary Fig. 10), therefore fed-batch fermentation was performed at 30 °C. Fed-batch culture of the GAS3 strain harbouring pTacCer1FadD and pTrcAcrTesA(L109P) at 30 °C resulted in the production of 580.8 mg l⁻¹ of hydrocarbons composed of 327.8 mg l⁻¹ nonane, 136.5 mg l⁻¹ dodecane, 64.8 mg l⁻¹ tridecane, 42.8 mg l⁻¹ 2-methyl-dodecane and 8.9 mg l⁻¹ tetradecane (Fig. 3b, Supplementary Figs 6 and 7). Because *fadD*-deleted GAS2 harbouring 'TesA(L109P) produced C10 FFA dominantly, decarbonylation yielded mainly nonane (Fig. 3).

The *A. thaliana* fatty aldehyde decarbonylase (CER1) is known to be active towards long-chain fatty aldehydes (>30 carbons)²⁶. However, our results suggest that this enzyme is also active towards short-chain fatty aldehydes (Supplementary Table 2). Furthermore, short-chain fatty alcohols as well as trace amounts of fatty aldehydes were also detected at the end of fermentation. Fatty alcohols were probably produced owing

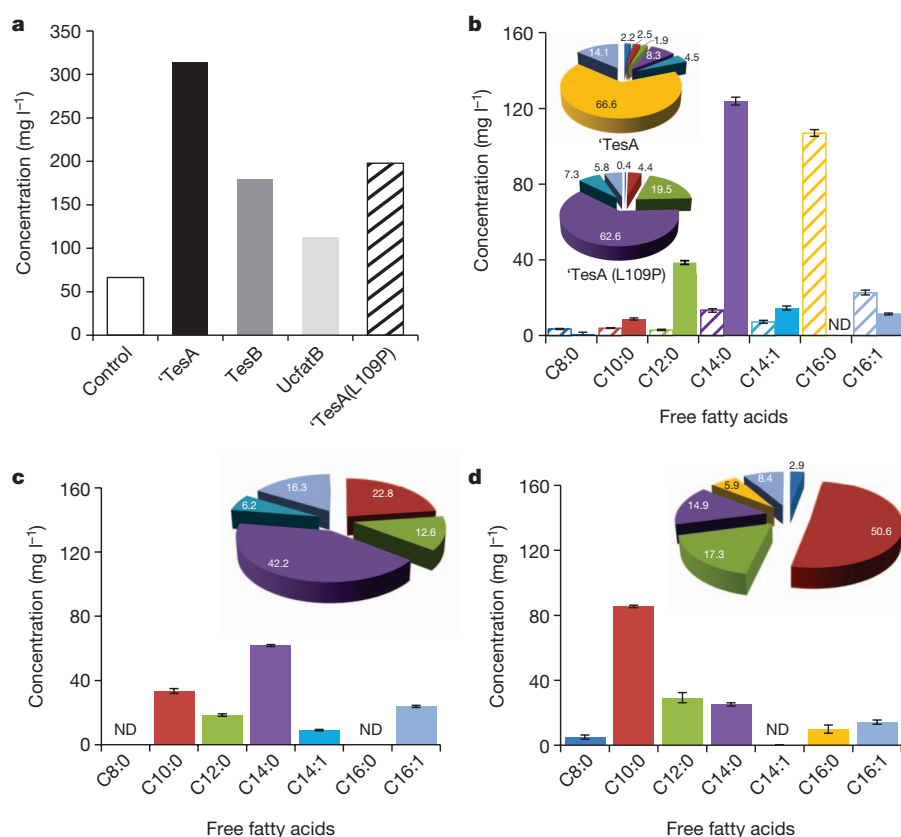


Figure 2 | Effects of three types of acyl-ACP thioesterases and 'TesA(L109P) on free fatty acid production. **a**, The total amounts of FFAs produced by the overexpression of 'TesA, TesB, UcfatB and 'TesA(L109P) in the *fadD*-deleted W3110 strain together with the control strain harbouring an empty vector. **b**, Distribution of FFAs produced in the *fadD*-deleted W3110 strain overexpressing 'TesA (hatched bar) and 'TesA(L109P) (solid bar). The percentage ratios of FFAs produced are also shown. **c**, **d**, The chain length distribution and percentage ratios of FFAs produced in *fadD*-deleted GAS1 (**c**) and *fadD*-deleted GAS2 (**d**) expressing 'TesA(L109P) are shown. Error bars represent the s.d. of experiments conducted in triplicate. Enzymes shown are: UcfatB, thioesterase of *Umbellularia californica*; TesB, thioesterase II of *E. coli*; 'TesA, the leaderless thioesterase I of *E. coli*; and 'TesA(L109P), a mutated leaderless thioesterase I of *E. coli*. ND, not detected.

to the presence of inherent *E. coli* alcohol dehydrogenase, despite its low activity under aerobic conditions, and also due to the overexpression of fatty acyl-CoA reductase, which has been shown to convert fatty aldehydes to fatty alcohols^{9,27}. Thus, the activity of aldehyde decarboxylase needs to be improved to further enhance the production of SCAs (see Supplementary Discussion). One possible way might be the overexpression of CER3, which was recently discovered to enhance the activity of CER1 (refs 28, 29).

Recently an *E. coli* strain integrated with a dynamic sensor-regulator system was developed for the production of C12–C20 long-chain fatty ethyl esters (FAEEs)⁵. Using GAS3 as a platform strain, production

of short-chain FAEEs was attempted. For the aerobic production of ethanol, the *E. coli* *adhE*^{mut} gene⁶ was expressed. For the esterification of fatty acyl-CoAs with ethanol, the *Acinetobacter* sp. ADP1 wax ester synthase (*atfA*) gene⁵ was expressed. Batch culture of the GAS3 strain harbouring the plasmids pTacAdhE^{mut}FadD and pTrcAtfA'TesA(L109P) allowed production of 477.7 mg l⁻¹ of short-chain FAEEs which consisted of 22.4 mg l⁻¹ C10, 363.1 mg l⁻¹ C12 and 92.2 mg l⁻¹ C14 FAEEs (Supplementary Fig. 8). Thus, the platform strain developed here can be used for the production of short-chain FAEEs as well as petrol by using different final metabolic pathways. However, it was interesting to note that the composition of FAEEs produced was somewhat different from

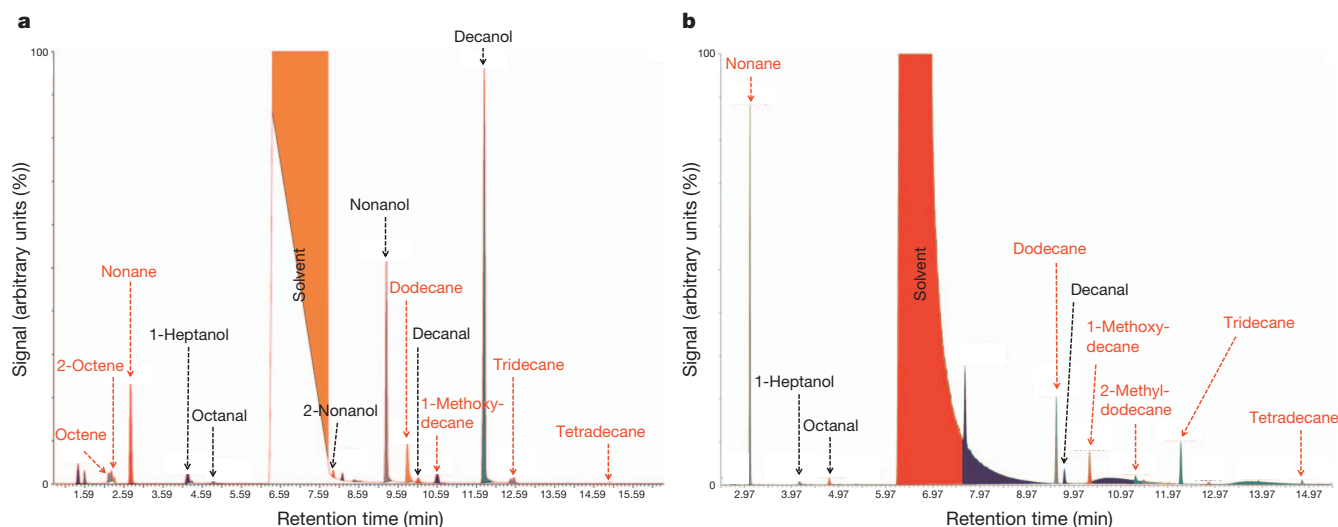


Figure 3 | GC-MS profile of fermentation products. The gas chromatography–mass spectrometry (GC-MS) profile of the hydrocarbon products obtained by fed-batch culture of the engineered GAS3 strain

harbouring pTacCer1FadD and pTrcAcR'TesA(L109P) at 31 °C (**a**) and 30 °C (**b**) are shown. The ion spectra of these compounds are shown in Supplementary Figs 4–7.

what would be expected from the results of SCAs production. As nonane was the major alkane produced, the expectation was to see C10 FAEE as the most abundant one rather than C12 FAEE. This discrepancy seems to be because of the substrate specificity of wax ester synthase, which has higher affinity towards long-chain fatty acyl-CoAs¹⁷.

We have developed new platform *E. coli* strains, the GAS2 strain capable of producing short-chain FFAs, and the GAS3 strain capable of producing SCHCs (petrol), fatty esters and fatty alcohols. This was possible by establishing the corresponding metabolic pathways by engineering *E. coli* fatty acid biosynthesis and degradation pathways and employing an engineered thioesterase. Also, if desired, long-chain alkanes suitable for diesel can be produced by employing the same platform strain together with non-engineered thioesterase (Supplementary Fig. 9). This work will serve as a stepping stone for establishing bioprocesses for the production of short-chain fatty acid derived chemicals and fuels from renewable resources.

METHODS SUMMARY

Bacterial strains and plasmids. *E. coli* strains, plasmids and oligonucleotides used are listed in Supplementary Tables 3 and 4. Detailed procedures for the construction of strains are described in the Methods. All DNA manipulations were performed according to standard procedures³⁰. All oligonucleotides were synthesized at GenoTech or Bioneer (Daejeon). Preparation of plasmids and DNA fragments was performed with Qiagen kits. All other chemicals used were of analytical grade and purchased from Sigma-Aldrich.

Culture condition and analysis. Recombinant *E. coli* strains were grown in MR medium (pH 6.8; see Methods) containing 10 g l⁻¹ glucose and 3 g l⁻¹ yeast extract at 31 °C and shaking at 220 r.p.m. Hydrocarbons were identified and quantified by gas chromatography-mass spectrometry (GC-MS) (Perkin Elmer Turbo Mass Clarus 600 coupled with a quadrupole mass selective detector on EI operated at 70 eV; see Methods for details). Retention times and fragmentation patterns were compared with GC-MS library database (NIST MS Search 2.0).

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 March; accepted 8 August 2013.

Published online 29 September 2013.

- Han, L., Lobo, S. & Reynolds, K. A. Characterization of β -ketoacyl-acyl carrier protein synthase III from *Streptomyces glaucescens* and its role in initiation of fatty acid biosynthesis. *J. Bacteriol.* **180**, 4481–4486 (1998).
- Heath, R. J. & Rock, C. O. Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. *J. Biol. Chem.* **271**, 1833–1836 (1996).
- Nunn, W. D., Giffin, K., Clark, D. & Cronan, J. E. Jr. Role for *fadR* in unsaturated fatty acid biosynthesis in *Escherichia coli*. *J. Bacteriol.* **154**, 554–560 (1983).
- Lo, Y. C., Lin, S. C., Shaw, J. F. & Liaw, Y. C. Substrate specificities of *Escherichia coli* thioesterase I/protease I/lysophospholipase L1 are governed by its switch loop movement. *Biochemistry* **44**, 1971–1979 (2005).
- Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnol.* **30**, 354–359 (2012).
- Holland-Staley, C. A., Lee, K., Clark, D. P. & Cunningham, P. R. Aerobic activity of *Escherichia coli* alcohol dehydrogenase is determined by a single amino acid. *J. Bacteriol.* **182**, 6049–6054 (2000).
- Peralta-Yahya, P. P., Zhang, F., del Cardayre, S. B. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320–328 (2012).
- Lennen, R. M., Braden, D. J., West, R. A., Dumesic, J. A. & Pfleger, B. F. A process for microbial hydrocarbon synthesis: overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnol. Bioeng.* **106**, 193–202 (2010).
- Schirmer, A., Rude, M. A., Li, X., Popova, E. & del Cardayre, S. B. Microbial biosynthesis of alkanes. *Science* **329**, 559–562 (2010).
- Harger, M. *et al.* Expanding the product profile of a microbial alkane biosynthetic pathway. *ACS Synthet. Biol.* **2**, 59–62 (2013).
- Howard, T. P. *et al.* Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **110**, 7636–7641 (2013).
- Altin, O. & Eser, S. Carbon deposit formation from thermal stressing of petroleum fuels. *Am. Chem. Soc. Div. Fuel Chem.* **49**, 764–766 (2004).
- Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86–89 (2008).
- Choi, Y. J., Park, J. H., Kim, T. Y. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for the production of 1-propanol. *Metab. Eng.* **14**, 477–486 (2012).
- Gary, J. H. & Handwerk, G. E. *Petroleum Refining: Technology and Economics* 4th edn (Marcel Dekker, 2001).
- Naggert, J. *et al.* Cloning, sequencing, and characterization of *Escherichia coli* thioesterase II. *J. Biol. Chem.* **266**, 11044–11050 (1991).
- Steen, E. J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559–562 (2010).
- Pollard, M. R., Anderson, L., Fan, C., Hawkins, D. J. & Davies, H. M. A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of *Umbellularia californica*. *Arch. Biochem. Biophys.* **284**, 306–312 (1991).
- Jing, F. *et al.* Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochem.* **12**, 44 (2011).
- Zheng, Y. *et al.* Boosting the free fatty acid synthesis of *Escherichia coli* by expression of a cytosolic *Acinetobacter baylyi* thioesterase. *Biotechnol. Biofuels* **5**, 76 (2012).
- Torella, J. P. Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proc. Natl. Acad. Sci. USA* **110**, 11290–11295 (2013).
- Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. Isolation and characterization of the β -ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J. Biol. Chem.* **267**, 6807–6814 (1992).
- Na, D. *et al.* Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nature Biotechnol.* **31**, 170–174 (2013).
- Zhang, H., Wang, P. & Qi, Q. Molecular effect of FadD on the regulation and metabolism of fatty acid in *Escherichia coli*. *FEMS Microbiol. Lett.* **259**, 249–253 (2006).
- Aarts, M. G., Keijzer, C. J., Stiekema, W. J. & Pereira, A. Molecular characterization of the *CER1* gene of *Arabidopsis* involved in epicuticular wax biosynthesis and pollen fertility. *Plant Cell* **7**, 2115–2127 (1995).
- McNevin, J. P., Woodward, W., Hannoufa, A., Feldmann, K. A. & Lemieux, B. Isolation and characterization of *eceriferum* (*cer*) mutants induced by T-DNA insertions in *Arabidopsis thaliana*. *Genome* **36**, 610–618 (1993).
- Reiser, S. & Somerville, C. Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase. *J. Bacteriol.* **179**, 2969–2975 (1997).
- Bernard, A. *et al.* Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell* **24**, 3106–3118 (2012).
- Bourdenx, B. Overexpression of *Arabidopsis* ECERIFERUM1 promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stress. *Plant Physiol.* **156**, 29–45 (2011).
- Sambrook, J. R. D. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank Y. H. Lee for her assistance in cloning work and S. J. Choi for performing the fermentation experiments for checking reproducibility. This work was supported by the Advanced Biomass Research and Development Center of Korea (ABC-2010-0029799) through the Global Frontier Research Program of the Ministry of Science, ICT and Future Planning (MSIP) through the National Research Foundation (NRF). Systems metabolic engineering work was supported by the Technology Development Program to Solve Climate Changes on Systems Metabolic Engineering for Biorefineries (NRF-2012-C1AAA001-2012M1A2A2026556) by MSIP through NRF.

Author Contributions S.Y.L. conceived and supervised the project. Y.J.C. performed all experiments and analysed the data. Y.J.C. and S.Y.L. wrote the manuscript together. Both authors approved the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.Y.L. (leesy@kaist.ac.kr).

Adrenaline-activated structure of β_2 -adrenoceptor stabilized by an engineered nanobody

Aaron M. Ring^{1,2*}, Aashish Manglik^{1*}, Andrew C. Kruse^{1*}, Michael D. Enos^{1,2}, William I. Weis^{1,2}, K. Christopher Garcia^{1,2,3} & Brian K. Kobilka¹

G-protein-coupled receptors (GPCRs) are integral membrane proteins that have an essential role in human physiology, yet the molecular processes through which they bind to their endogenous agonists and activate effector proteins remain poorly understood. So far, it has not been possible to capture an active-state GPCR bound to its native neurotransmitter. Crystal structures of agonist-bound GPCRs have relied on the use of either exceptionally high-affinity agonists^{1,2} or receptor stabilization by mutagenesis^{3–5}. Many natural agonists such as adrenaline, which activates the β_2 -adrenoceptor (β_2 AR), bind with relatively low affinity, and they are often chemically unstable. Using directed evolution, we engineered a high-affinity camelid antibody fragment that stabilizes the active state of the β_2 AR, and used this to obtain crystal structures of the activated receptor bound to multiple ligands. Here we present structures of the active-state human β_2 AR bound to three chemically distinct agonists: the ultrahigh-affinity agonist BI167107, the high-affinity catecholamine agonist hydroxybenzyl isoproterenol, and the low-affinity endogenous agonist adrenaline. The crystal structures reveal a highly conserved overall ligand recognition and activation mode despite diverse ligand chemical structures and affinities that range from 100 nM to ~80 pM. Overall, the adrenaline-bound receptor structure is similar to the others, but it has substantial rearrangements in extracellular loop three and the extracellular tip of transmembrane helix 6. These structures also reveal a water-mediated hydrogen bond between two conserved tyrosines, which appears to stabilize the active state of the β_2 AR and related GPCRs.

GPCRs relay extracellular signals across a cell membrane by means of a conformational change after the binding of an extracellular agonist. GPCR activation by endogenous agonists remains poorly understood owing to the paucity of active receptor structures that have been elucidated in complex with agonists. Although a number of GPCRs have been crystallized in recent years, only the β_2 AR and rhodopsin have been crystallized in fully active states^{1,6}, and in both cases structures are available only for complexes with a single agonist. Owing to the conformational plasticity and biochemical instability of agonist-bound receptors⁷, the few agonist-bound structures of GPCRs solved thus far have relied on the use of covalent¹ or extremely high-affinity agonists², crystallographic chaperones to trap active states (a G protein⁸ or antibody fragment⁹), or thermostabilizing mutations³. The last approach has only yielded structures of agonist-occupied receptor in partially active^{3,4} or inactive⁵ conformations.

To understand better how diverse agonists can activate a single receptor, we developed a strategy for stabilizing active-state structures of the β_2 AR bound to low-affinity agonists including the natural agonist adrenaline. Here, we describe the directed evolution of Nb80, a conformationally selective single-domain camelid antibody fragment (nanobody) that was used to obtain the first active-state structure of the β_2 AR⁹. Comparison with the structure of the β_2 AR in complex with the G protein G_s confirmed that Nb80 stabilizes a physiologically relevant active state⁸. However, the β_2 AR–Nb80 structure was of modest resolution

(3.5 Å) and crystals could only be obtained with the high-affinity agonist BI167107; crystallization trials with catecholamine agonists were unsuccessful despite extensive screening. We reasoned that improving the affinity of Nb80 for agonist-bound β_2 AR would decrease receptor conformational heterogeneity and enable crystallization of the receptor bound to low-affinity agonists. However, directed evolution of conformationally selective GPCR-binding proteins has never been described, probably owing to the challenges involved in biochemical manipulation of integral membrane proteins. We used yeast surface display together with a conformationally specific selection strategy to improve the binding affinity of Nb80 while maintaining its conformational selectivity. The resulting high-affinity variants retain their specificity for the active state of the receptor, which was characteristic of the original Nb80. Using the high-affinity variant Nb6B9, we determined a high-resolution (2.8 Å) active-state structure of the β_2 AR bound to BI167107, and also determined the structures of the β_2 AR bound to two catechol-containing agonists: hydroxybenzyl isoproterenol (HBI) and adrenaline, an endogenous low-affinity agonist of the β_2 AR, at 3.1 Å and 3.2 Å resolution, respectively.

To assess the feasibility of engineering Nb80, we displayed Nb80 on the surface of the yeast strain EBY100 as an amino-terminal fusion to the yeast cell-wall protein Aga2p (Fig. 1a). Yeast displaying Nb80 were stained with purified, detergent-solubilized, biotinylated β_2 AR after pre-incubation of receptor with the agonist BI167107 or the inverse agonist carazolol. Nb80-displaying yeast specifically bound to β_2 AR with an overwhelming preference for agonist-occupied receptor (Fig. 1b), with a half-maximum effective concentration (EC₅₀) of 140 nM (Supplementary Fig. 1a). Next, we constructed a library of Nb80 mutants in which residues at the receptor-binding surface were randomized with conservative substitutions (Supplementary Fig. 2). The library was subjected to six rounds of selection (Fig. 1c). First, the library was positively selected with decreasing concentrations of BI167107-bound β_2 AR. Before positive selection in rounds 2–5, the library was negatively selected against binding to inverse-agonist-occupied β_2 AR in order to remove variants that had lost conformational specificity. For the final round of selection, we enriched variants with the slowest dissociation rates. Receptor rebinding was blocked by the addition of a large excess of soluble Nb80 after the initial receptor-binding step (Supplementary Fig. 3). This selection strategy resulted in a progressive increase in binding affinity for agonist-occupied receptor without a similar increase in binding to inverse-agonist-occupied receptor (Fig. 1d).

Nanobody 6B9 (Nb6B9) was chosen from 23 variants screened from the final round of selection (Supplementary Fig. 4) as it represented one of the highest-affinity binders tested, contained mutations that reached consensus among all sequenced clones, and was the most prevalent sequence observed. We expressed and purified Nb6B9 and Nb80, and then used surface plasmon resonance (SPR) to measure binding kinetics and affinities. Nb6B9 bound to BI167107-occupied

¹Department of Molecular and Cellular Physiology, Stanford University, Stanford, California 94305, USA. ²Department of Structural Biology, Stanford University, Stanford, California 94305, USA. ³Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA.

*These authors contributed equally to this work.

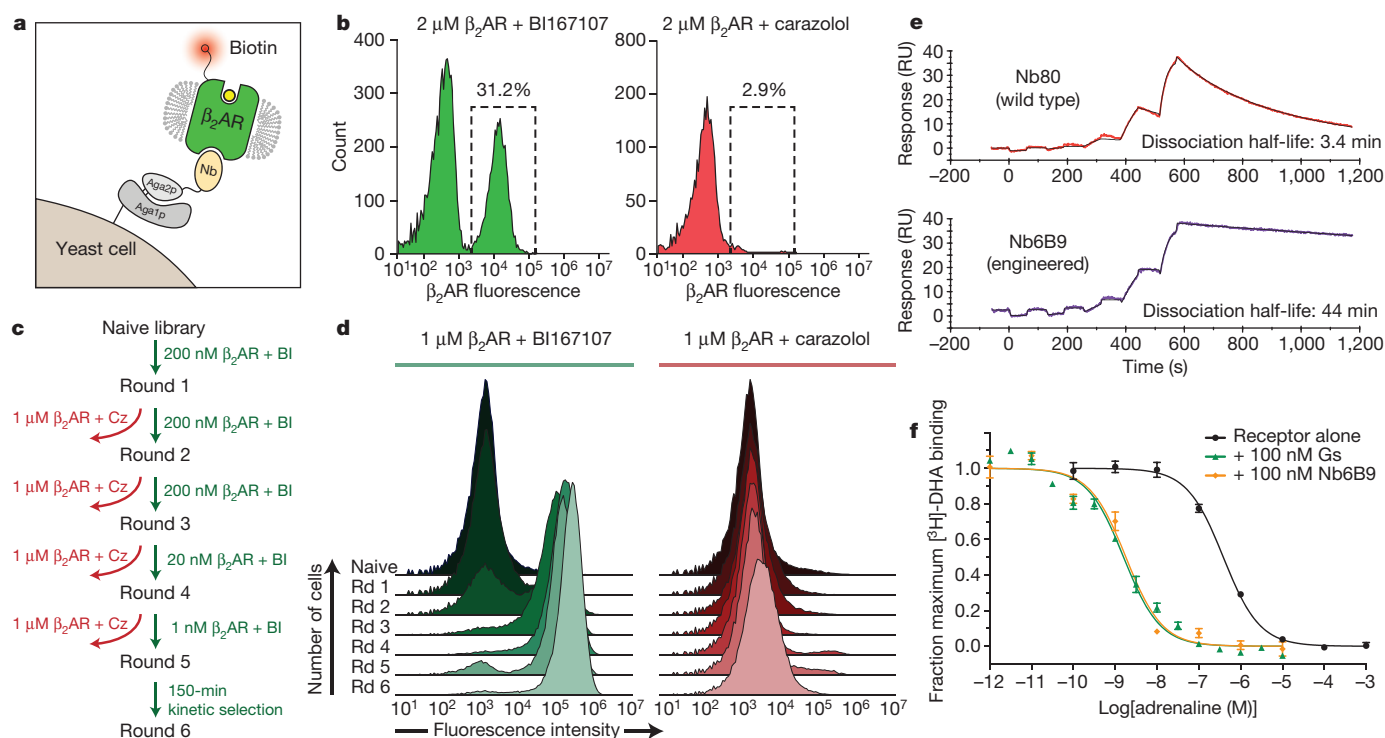


Figure 1 | Conformational selection of nanobodies and characterization of high-affinity Nb6B9. **a**, Schematic representation of yeast display of Nb80. Nb80 is fused to the N terminus of Aga2p, which attaches to the yeast cell wall through a covalent interaction with Aga1p. **b**, Staining of Nb80-expressing yeast with β_2 AR bound to the agonist BI167107 (left) or the inverse agonist carazolol (right). Per cent of yeast within the boxed gate is indicated. **c**, Flowchart summary of conformational selection process. BI, BI167107; Cz, carazolol. **d**, Histogram overlays assessing β_2 AR staining of the library at each round (Rd) of selection. The left panel shows staining with 1 μ M BI167107-occupied

receptor, and the right panel shows staining with 1 μ M carazolol-occupied receptor. **e**, Representative single-cycle kinetics SPR sensorgram of wild-type Nb80 (top) and engineered Nb6B9 (bottom) binding immobilized β_2 AR bound to BI167107. RU, response unit. **f**, 3 H-dihydroalprenolol (3 H-DHA) competition binding shows a comparable increase in β_2 AR affinity for adrenaline in the presence of Nb6B9 as with G protein G_s . 3 H-DHA affinity is largely unchanged in the presence of Nb6B9 (Supplementary Table 2). Data and error bars represent the mean \pm standard error of the mean from three experiments.

β_2 AR with an affinity of 6.4 nM, a near tenfold improvement over Nb80 (Fig. 1e). This increase in affinity resulted from a 13-fold reduction in the dissociation rate. Competition binding experiments revealed that the

β_2 AR bound adrenaline with a high affinity in the presence of 100 nM Nb6B9, which is comparable to the affinity observed in the presence of the G protein G_s (Fig. 1f).

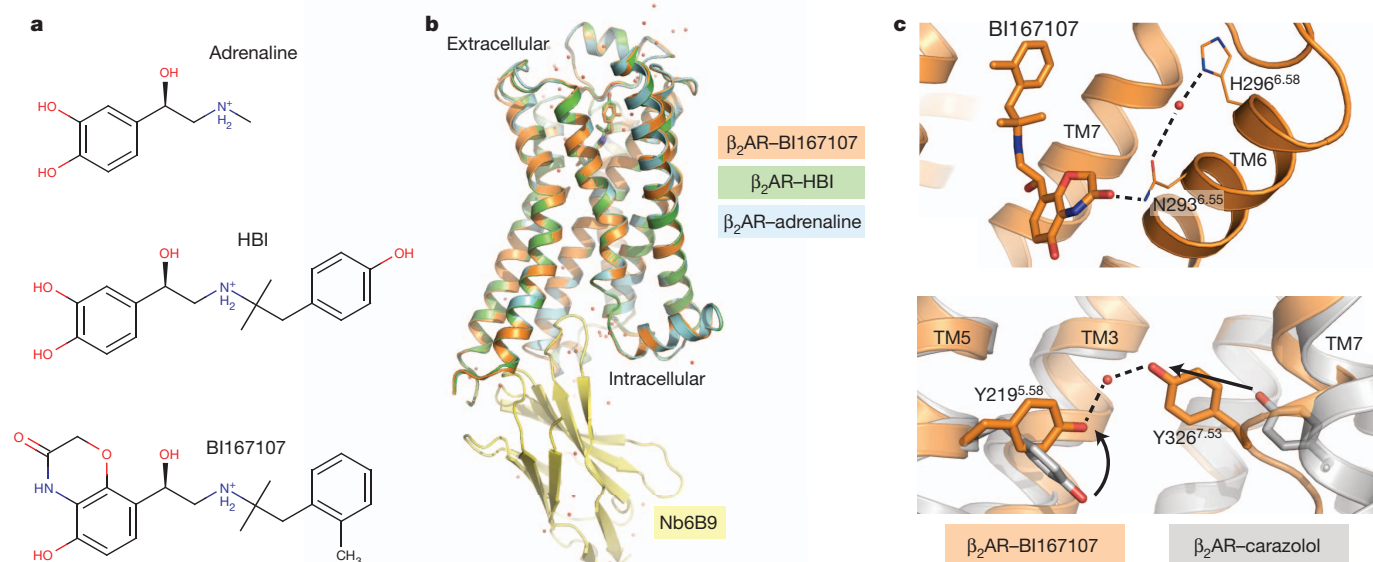


Figure 2 | Structure of the activated β_2 AR in complex with three agonists. **a**, Chemical structures of the three ligands used for crystallization trials. **b**, All three active-state structures, showing remarkable similarity in overall receptor conformation. **c**, The 2.8 Å resolution structure of BI167107-bound β_2 AR reveals active-state water molecules: a bridging water molecule participates in a

polar network at the ligand-binding site (top) and a second water molecule mediates a hydrogen bond between two highly conserved tyrosines. Such an interaction is possible in the active state (orange) but not the inactive state (grey).

We used the lipidic mesophase method¹⁰ to crystallize complexes of Nb6B9 with β_2 AR bound to three different ligands, shown in Fig. 2a. Crystals grown with β_2 AR bound to the high-affinity agonist BI167107 showed strong diffraction, and a structure was obtained to 2.8 Å resolution (Supplementary Table 1). This represents a significant improvement over the previous 3.5 Å structure of β_2 AR bound to the same ligand⁹. This higher-resolution structure showed few differences from the Nb80 complex structure (Supplementary Fig. 5). Mutations in Nb6B9 appear to increase shape complementarity to active β_2 AR (Supplementary Fig. 6). Many water molecules were clearly resolved for the first time, particularly in the extracellular region of the receptor (Fig. 2b, c). On the intracellular side of the receptor, a water molecule was found to mediate a hydrogen bond between Tyr 326^{7,53} of the NPxxY motif and the highly conserved Tyr 219^{5,58} on the intracellular side of transmembrane helix 5 (TM5), similar to a water seen in a recent structure of metarhodopsin II¹¹. Electron density suggestive of a water molecule was also seen in HBI- and adrenaline-bound β_2 AR structures, despite their slightly lower resolution. The water-mediated hydrogen bond between Tyr 219^{5,58} and Tyr 326^{7,53} is possible only in the active conformation of the receptor (Fig. 2c), and the observed water-mediated hydrogen bond may therefore contribute to active state stability in the β_2 AR and other GPCRs, serving as an active-state

counterpart to the ‘ionic lock’ that stabilizes the inactive state¹². In support of this notion, mutation of the corresponding Tyr 223^{5,58} to phenylalanine in rhodopsin decreases the stability of the meta II state¹³ and greatly reduces activation of transducin¹⁴. Moreover, mutation of Tyr 227^{5,58} to alanine resulted in the largest increase in thermostability for the inactive-state thermostabilized β_1 AR¹⁵.

Although BI167107 exhibits many features typical of β_2 AR agonists, it lacks the catechol moiety of the endogenous agonists adrenaline and noradrenaline. Hence, it is conceivable that these agonists stabilize a different conformation of the activated receptor-binding pocket. To assess this possibility, we pursued crystallographic studies of complexes of Nb6B9 with β_2 AR bound to the low-affinity endogenous agonist adrenaline and the high-affinity catecholamine agonist HBI. In each case, crystals could be grown in nearly identical conditions to those for the BI167107 complex, with clear electron density to identify the position and orientation of each ligand (Supplementary Fig. 7).

Despite the chemical diversity of these ligands, the structures of β_2 AR bound to the catecholamine agonists and to BI167107 have very similar overall structures (Fig. 3a, b). A notable exception is a shift in the position of Asn 293^{6,55}, which was previously determined to hydrogen bond with the amide carbonyl on the head group of BI167107. The smaller catechol ring of adrenaline and HBI precludes hydrogen

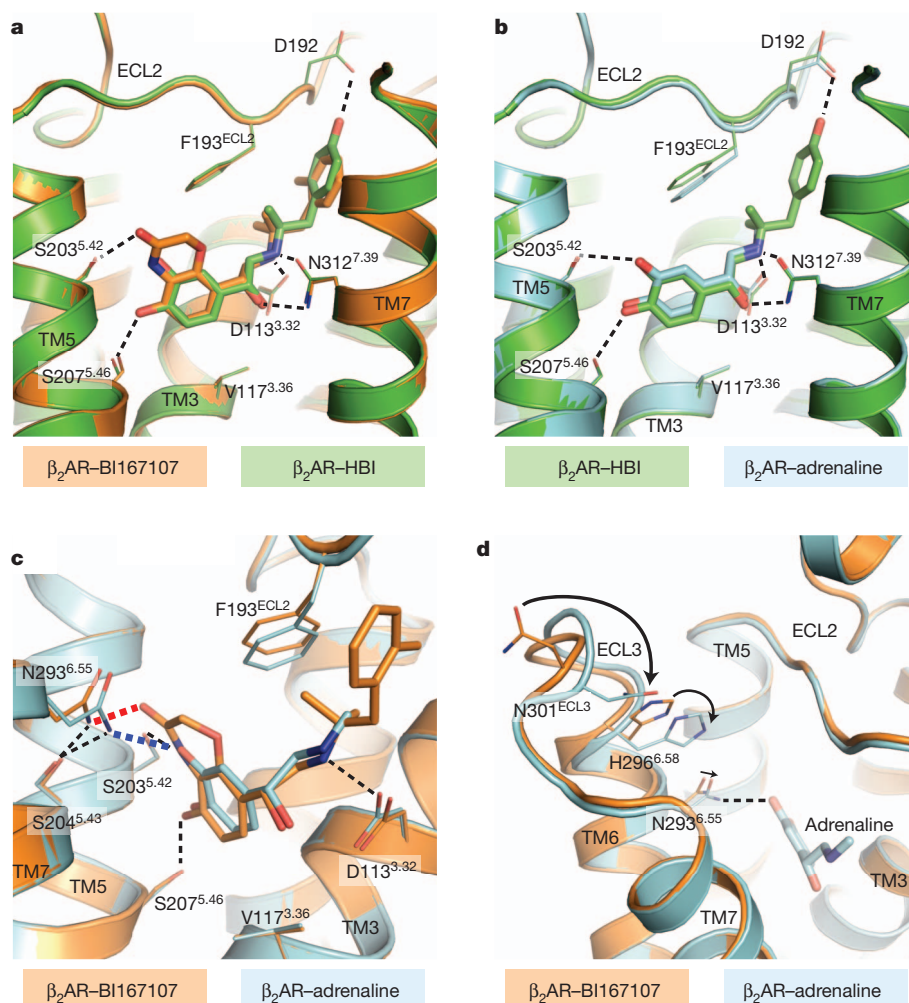


Figure 3 | Comparison of agonist-binding modes. **a**, Comparison of BI167107-bound receptor (orange) with HBI-bound receptor (green) shows a highly conserved agonist-binding mode. **b**, Similarly, adrenaline-bound (cyan) and HBI-bound (green) receptor structures are highly similar. **c**, An analogous comparison of BI167107-bound β_2 AR (orange) with adrenaline-bound receptor (cyan) shows the similar polar networks for the two ligands (black

dotted lines) with a notable difference in the hydrogen bonding of Asn 293^{6,55} to the amide proton in BI167107 (red dotted line) or the *meta* hydroxyl of adrenaline (blue dotted line). **d**, Owing to this difference, Asn 293^{6,55} and TM6 shift inwards in the adrenaline-bound structure, leading to a cascade of changes culminating in a rearrangement of ECL3.

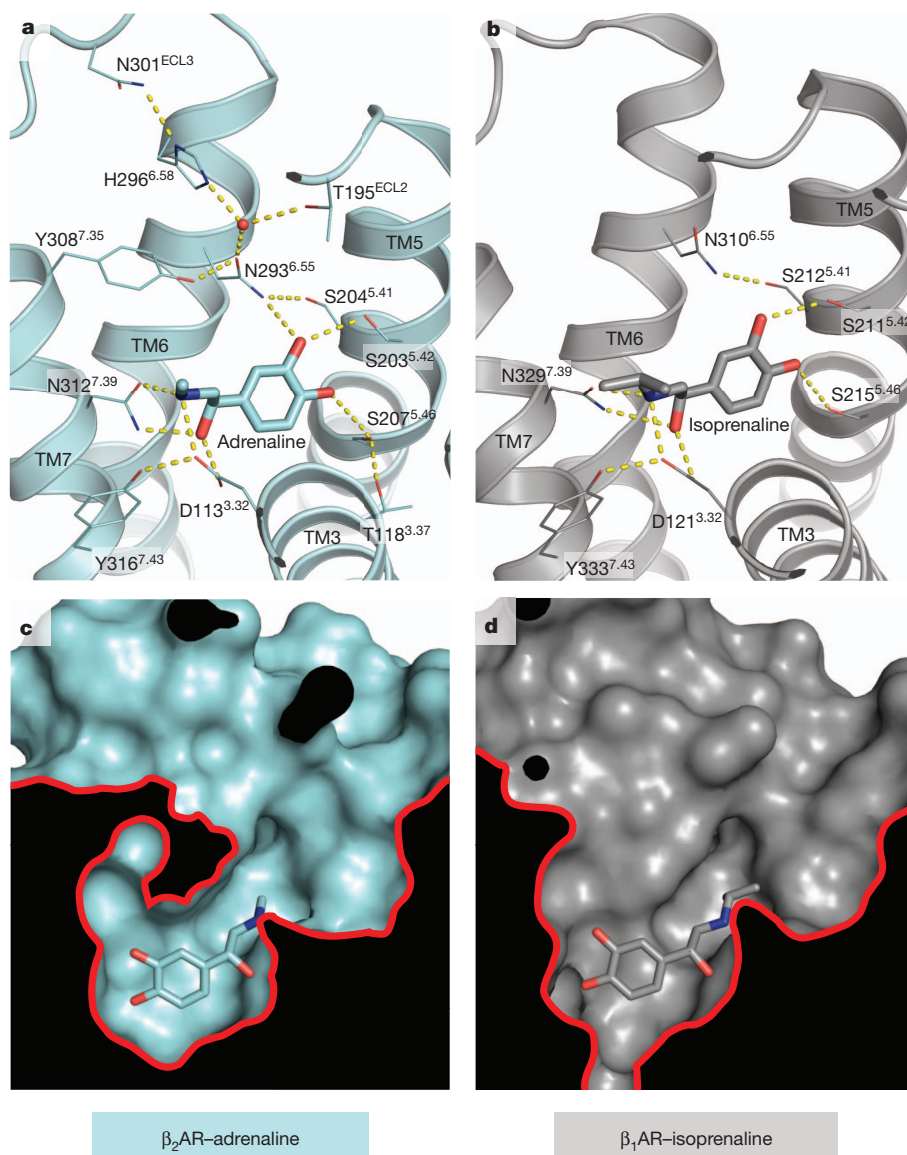


Figure 4 | Activation by catecholamine agonists. a–d, For the first time, structures of catecholamine-bound adrenergic receptors in active and inactive conformations can be compared. a, The structure of β_2 AR in an active conformation bound to the agonist adrenaline reveals an extended polar contact network linking the orthosteric site to ECL2 and 3, whereas the

bonding with Asn 293^{6.55} in the receptor conformation observed in the BI167107-bound structure. To maintain the corresponding hydrogen bond between Asn 293 and the *meta* hydroxyl moiety on the catechol ring, the receptor undergoes a 1.2 Å shift in the extracellular side of TM6, which bends towards the ligand (Fig. 3c). This shift alters the hydrogen-bonding network in this region and thereby causes a change in the conformation of His 296^{6.58}. For adrenaline-bound β_2 AR, the TM6 conformational change is further propagated towards the extracellular side of the receptor, leading to a conformational rearrangement in extracellular loop 3 (ECL3; Fig. 3d). This change also alters the extracellular surface of the receptor, with adrenaline-bound β_2 AR having a contracted extracellular vestibule (Supplementary Fig. 8).

The relatively subtle differences in receptor conformation observed for the different co-crystallized agonists suggest that the activation mechanism of the β_2 AR is highly similar for all agonists. Much like BI167107, the catechol head groups of adrenaline and HBI engage β_2 AR residues previously characterized to be important for agonist binding and receptor activation (Supplementary Fig. 9). Consistent with prior mutagenesis studies, Ser 203^{5.42} and Ser 207^{5.46} make hydrogen

structure of thermostabilized β_1 AR in an inactive conformation bound to a similar catecholamine agonist, isoprenaline, shows a far more limited polar network. c, Likewise, a surface view of the active-state structure (c) shows a substantial contraction of the binding site compared with the inactive β_1 AR structure (d).

bonds with the catecholamine phenoxy moieties¹⁶ (Fig. 3a–c), and the conformation of these residues is nearly identical to that observed for the BI167107 head group. Ser 204^{5.43}, which was previously thought to contact the para-hydroxy group of the catecholamine directly^{16,17}, engages the catecholamine head group indirectly in an extended polar network with Tyr 308^{7.35} and Asn 293^{6.55}. However, unlike BI167107-bound β_2 AR, this polar interaction network is extended by inclusion of His 296^{6.58} in the catecholamine-bound receptor, suggesting that conformational rearrangement of His 296^{6.58} may stabilize the slightly smaller orthosteric binding pocket observed for catecholamine agonists. The agonist-induced rearrangements in the central portion of the transmembrane segments and intracellular surface are virtually identical in all three agonist-bound structures (Fig. 3b and Supplementary Fig. 10). Structures of β_2 AR bound to BI167107 and the catecholamine agonists all show very similar activation-related changes in the residues that connect the orthosteric ligand-binding pocket to the intracellular surface, suggesting that the mechanism for allosteric coupling between the orthosteric binding site and the G-protein-coupling domain is probably a conserved feature of β_2 AR activation. Therefore, different

agonists stabilize the same conformational rearrangements in the receptor through different chemical interactions.

In contrast to the marked similarity in receptor conformation for all three agonists crystallized here, far more substantial conformational differences are seen relative to a previously reported structure of the thermostabilized turkey β_1 -adrenoceptor (β_1 AR) bound to the catecholamine agonist isoproterenol⁵. Probably due to the thermostabilization procedure, the overall receptor conformation of isoprenaline-bound β_1 AR closely resembles that of the antagonist-bound, inactive β_1 AR¹⁸, as well as that of covalent-agonist-bound, inactive β_2 AR¹ (Supplementary Fig. 11). Thus, a comparison of the structural changes between inactive β_1 AR and active β_2 AR, each bound to catecholamines, offers new insight into how agonists bind adrenoceptors both in the low-affinity state and the high-affinity, G-protein-coupled state. Within the binding pocket, isoprenaline makes a hydrogen bond with Ser 211^{5,42} and the β -hydroxylamine moiety engages conserved residues Asp^{3,32} and Asn^{7,39} in a very similar manner to the active-state structures (Fig. 4a, b). Similar interactions can occur in both active and inactive states, probably accounting for the fact that the β -hydroxylamine moiety is an important feature of both β -adrenoceptor agonists and antagonists/inverse agonists. However, the isoprenaline catechol head group engages a limited network of polar contacts in the inactive β_1 AR structure, whereas adrenaline bound to active β_2 AR engages an extensive polar network linking the orthosteric site to the extracellular loops (Fig. 4a). As a consequence of structural changes stabilized by this polar network, the catechol head group of adrenaline is nearly completely enclosed within the orthosteric binding pocket of activated β_2 AR (Fig. 4c). In comparison, isoprenaline bound to inactive β_1 AR is highly exposed to the extracellular solvent and is slightly displaced towards the extracellular side of the receptor (Fig. 4d). Thus, in β -adrenoceptors, the combination of a more extensive polar network and a smaller binding pocket probably accounts for the enhanced agonist affinity seen in the presence of either G_s or G protein mimetic nanobodies. Moreover, such differences between active and inactive structures highlight the importance of active-state GPCR crystal structures in understanding the structural basis for agonist activity.

In conclusion, the use of new approaches in combinatorial biology has led to the development of Nb6B9, an exceptionally high-affinity GPCR-stabilizing nanobody. This molecule exhibits enhanced affinity for β_2 AR relative to wild-type Nb80, and it enabled the crystallization of β_2 AR in complex with three different agonists with diverse chemical structures and a wide range of affinities. The use of such high-affinity crystallization chaperones may be generally useful in the determination of active-state structures of GPCRs bound to low-affinity agonists. The crystallographic studies presented here reveal subtle, ligand-specific differences in receptor conformation superimposed on the backdrop of an overall conserved agonist-binding mode and activation mechanism, offering new insight into how chemically diverse agonists can activate a single receptor.

METHODS SUMMARY

Nb80 was displayed on the surface of EBY100 yeast as an N-terminal fusion to Aga2p, and an affinity-maturation library was generated by assembly PCR. Yeast were stained with detergent-solubilized receptor, and selections were carried out using magnetic-activated cell sorting. For crystallography, the β_2 AR with an N-terminal T4 lysozyme (T4L) fusion was expressed in Sf9 insect cells and purified by ligand-affinity chromatography. Nb80 and Nb6B9 were expressed in the *Escherichia coli* periplasm and purified by Ni-NTA chromatography. For crystallization, T4L- β_2 AR was incubated with ligand, followed by the addition of excess Nb6B9 and purification by gel filtration. The purified complex was reconstituted into the lipidic cubic phase and crystallized. Diffraction data were collected at Advanced Photon Source

GM/CA beamlines 23ID-B and 23ID-D, and the structures were solved by molecular replacement.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 May; accepted 13 August 2013.

Published online 22 September 2013.

- Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist- β_2 adrenoceptor complex. *Nature* **469**, 236–240 (2011).
- Xu, F. *et al.* Structure of an agonist-bound human A_{2A} adenosine receptor. *Science* **332**, 322–327 (2011).
- White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
- Lebon, G. *et al.* Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature* **474**, 521–525 (2011).
- Warne, T. *et al.* The structural basis for agonist and partial agonist action on a β_1 -adrenergic receptor. *Nature* **469**, 241–244 (2011).
- Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).
- Nygaard, R. *et al.* The dynamic process of β_2 -adrenergic receptor activation. *Cell* **152**, 532–542 (2013).
- Rasmussen, S. G. *et al.* Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).
- Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175–180 (2011).
- Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols* **4**, 706–731 (2009).
- Deupi, X. *et al.* Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc. Natl Acad. Sci. USA* **109**, 119–124 (2012).
- Schneider, E. H., Schnell, D., Strasser, A., Dove, S. & Seifert, R. Impact of the DRY motif and the missing “ionic lock” on constitutive activity and G-protein coupling of the human histamine H₄ receptor. *J. Pharmacol. Exp. Ther.* **333**, 382–392 (2010).
- Elgeti, M. *et al.* Conserved Tyr223^{5,58} plays different roles in the activation and G-protein interaction of rhodopsin. *J. Am. Chem. Soc.* **133**, 7159–7165 (2011).
- Goncalves, J. A. *et al.* Highly conserved tyrosine stabilizes the active state of rhodopsin. *Proc. Natl Acad. Sci. USA* **107**, 19861–19866 (2010).
- Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. Conformational thermostabilization of the β_1 -adrenergic receptor in a detergent-resistant form. *Proc. Natl Acad. Sci. USA* **105**, 877–882 (2008).
- Strader, C. D., Candelore, M. R., Hill, W. S., Sigal, I. S. & Dixon, R. A. Identification of two serine residues involved in agonist activation of the β -adrenergic receptor. *J. Biol. Chem.* **264**, 13572–13578 (1989).
- Liapakis, G. *et al.* The forgotten serine. A critical role for Ser-203^{5,42} in ligand binding to and activation of the β_2 -adrenergic receptor. *J. Biol. Chem.* **275**, 37779–37788 (2000).
- Warne, T. *et al.* Structure of a β_1 -adrenergic G-protein-coupled receptor. *Nature* **454**, 486–491 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Hilger for critical reading of the manuscript. We acknowledge support from the Stanford Medical Scientist Training Program (A.M. and A.M.R.), the American Heart Association (A.M.), the National Science Foundation (A.C.K.), the Ruth L. Kirschstein National Research Service Award (A.M.R.), National Institutes of Health grants NS02847123 and GM08311806 (B.K.K.), from the Mathers Foundation (B.K.K., W.I.W. and K.C.G.), and from the Howard Hughes Medical Institute (K.C.G.).

Author Contributions A.M.R. designed and performed yeast display staining and selection experiments, nanobody expression, purification and characterization on yeast and by SPR. A.M. and A.C.K. designed and performed receptor expression, purification, radioligand-binding experiments and crystallography experiments. M.D.E. performed crystallization experiments with adrenaline-bound β_2 AR under the supervision of A.M. and A.C.K. The manuscript was written by A.M.R., A.M. and A.C.K. W.I.W. supervised structure refinement. K.C.G. and B.K.K. supervised experiments, and B.K.K. supervised manuscript preparation.

Author Information Coordinates and structure factors for the β_2 AR-Nb6B9 complexes with BI167107, HBI and adrenaline ligands are deposited in the Protein Data Bank under accession codes 4LDE, 4LDL and 4LDO, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.C.G. (kcgarcia@stanford.edu) or B.K.K. (kobilka@stanford.edu).

CORRIGENDUM

doi:10.1038/nature12667

Corrigendum: APOBEC3B is an enzymatic source of mutation in breast cancer

Michael B. Burns, Lela Lackey, Michael A. Carpenter, Anurag Rathore, Allison M. Land, Brandon Leonard, Eric W. Refsland, Delshanee Kotandeniya, Natalia Tretyakova, Jason B. Nikas, Douglas Yee, Nuri A. Temiz, Duncan E. Donohue, Rebecca M. McDougle, William L. Brown, Emily K. Law & Reuben S. Harris

Nature 494, 366–370 (2013); doi:10.1038/nature11881

We reported a comparison of the DNA cytosine deamination context of APOBEC3B *in vitro* with the observed C-to-T mutation context in breast cancer (see Fig. 4c of the original Letter). We incorrectly stated in the Fig. 4c legend that the data represent all cytosines. However, this analysis focused on unmodified DNA cytosines and excluded all potentially methylatable 5' CpG motifs, which, being more prone to spontaneous deamination, might have skewed the analysis by providing a false positive signal; this was noted in the original Methods

Summary. It was drawn to our attention by Thanos Halazonetis that the Fig.4c legend could mislead other researchers if they were not to take this exclusion into account. We therefore provide an updated version of Fig. 4c (see Fig. 1 of this Corrigendum) with the full logo representation for the *in vitro* preferences of APOBEC3B and the C-to-T mutation context in breast cancer. This does not affect the central conclusion of our study that APOBEC3B is an enzymatic source of mutation in breast cancer.

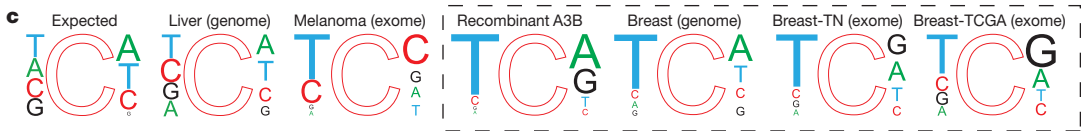


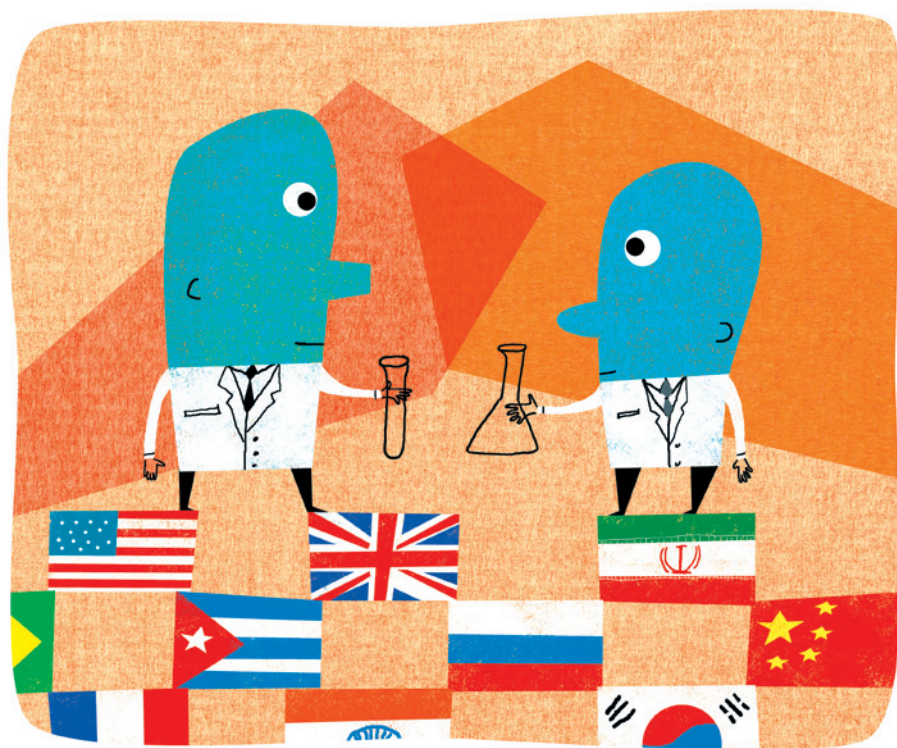
Figure 1 | This figure shows the corrected Fig. 4c of this Letter. c, Local sequence contexts for all genomic cytosines (expected), cytosines deaminated by recombinant A3B (Supplementary Fig. 13), and observed C-to-T transitions in the indicated cancers.

CAREERS

COLUMN Success entails acknowledging both achievements and potential **p.583**

NATUREJOBS FACEBOOK Science-careers advice and information go.nature.com/4lzxa

NATUREJOBS For the latest career listings and advice www.naturejobs.com



PROFESSIONAL SOCIETIES

Come together

Scientific organizations can help researchers — especially in developing countries — to make contacts and boost their skills.

BY KAREN KAPLAN

Environmental scientist Henry Roman was partway through his second postdoc when he stumbled across the website of the World Association of Young Scientists (WAYS) in 2006. He read about how the then-fledgling group was developing a global network to help early-career researchers to exchange job and careers information, form collaborations and promote their work. Intrigued, he registered — then promptly forgot about the whole thing.

But a few months later, Roman got an e-mail from a regional office of the International Council for Science in Pretoria. WAYS wanted to establish a branch in Africa and was meeting that year in Pretoria, where Roman

was working at the South African Council for Scientific and Industrial Research. Leaders wanted to invite him to the meeting; he accepted.

Today, Roman is chairman of WAYS-Africa, a task he fits around his busy job as director of environmental services and technologies for the Department of Science and Technology in Pretoria. He tackles the mission with passion: he sets the association's strategic direction, launches and manages partnerships, and raises funds. Roman says that his involvement in WAYS has helped him to build skills and develop his professional network in ways that he had not imagined as a student or early postdoc. He earned his degrees and completed his postdocs in his native South Africa, and before he joined WAYS he had rarely interacted with

anyone outside the country. But the association quickly introduced him to international colleagues. Through those contacts, he was invited to a 2009 international science forum in Budapest, where he met another environmental scientist from Pretoria — who turned out to be on the ministry's application-review panel when Roman applied for his current position. WAYS membership "has opened up the world to me", says Roman.

Early-career scientists have plenty of excuses not to join a scientific organization. Researchers spend hundreds of hours a week in the lab and have no spare time; they might have a partner, children or pets at home and be unable to travel to meetings; they might already be striving to build a network through social media and conferences (see 'Trouble maintaining members').

But joining and actively participating in associations for junior researchers can confer many advantages. It can help scientists to expand and grow their networks, which in turn can lead to new research ideas, collaborations, papers or even job offers. Members learn skills such as public speaking, fund-raising, organizing meetings, working in groups beyond the lab and navigating different cultures. And they can broaden their understanding of domestic and international science policy. Through WAYS, Roman says, he developed a perspective that vastly improved his chances of nabbing his current post. "It forces you to knock on many doors," he says.

MEMBERS WITH BENEFITS

The social pay-offs are obvious: members often form close friendships, and annual meetings typically feature parties that encourage mingling and having a drink or two. But scientific organizations also deliver more substantive returns. Some, among them the US National Postdoctoral Association (NPA) in Washington DC, focus on benefits such as providing advice on professional and career-development challenges and opportunities. Lorraine Tracey, chair of the board of directors of the NPA and a medical-science liaison at Teva Pharmaceuticals in Tampa, Florida, says that working with the NPA as a postdoc changed how she did science. "I was involved in conversations with people outside my lab and department about their research, and that informed mine," she says.

The Global Young Academy (GYA), based in Berlin, offers science-centred rewards. The three-year-old organization aims to bring together early-career researchers to ►

IMAGES.COM/CORBIS

► find remedies to global challenges such as tainted water and food supplies. Bernard Slippers, a GYA founding member, credits his previous work as academy co-chair and executive-committee member with helping him to reach career milestones: being named full professor in microbial ecology at the University of Pretoria, for example, and being invited to Thailand to attend a meeting of a joint programme of the US National Science Foundation (NSF) and the US Agency for International Development. He also credits his academy activity with helping him to launch a research project on new ecosystems and sustainability: he met his collaborators during a workshop co-presented by the GYA and two similar groups, the South African Young Academy of Science and the German Young Academy. “Being part of this global organization — what that has done for my perspective and output is profound,” says Slippers.

MEET AND GREET

Labs and conferences are often international, but they do not always allow for much interaction between disciplines; nor do they require researchers to take an active role in organizing or designing activities. Vinitha Thadhani, a member of the GYA and founding member and current president of the Sri Lankan Academy of Young Scientists in Colombo, recalls

the vast array of scientists that she met in China in 2008, during the GYA's pre-launch talks at an annual forum. Thadhani, a senior lecturer in chemistry at the University of Sri Jayewardenepura, was one of 43 scientists representing 32 nations at the forum. Through such events, “you come to know what other eminent young scientists around the globe do and [their] work in different fields”, she says.

When members discuss their research at meetings or on group chat boards, new ideas can emerge. In one case, Thadhani and the Sri Lankan academy were seeking expertise about the effects of a particular chemical on the environment and human health. They turned to the GYA, whose larger and broader membership includes chemists, economists, toxicologists and medical professionals. The GYA mentioned the issue in its newsletter,



“It’s provided opportunities to establish collaborations with people I would not meet in my regular line of work.”

Patrick Arthur

generating responses that contained enough information for the Sri Lankan academy to write an article that will be published on its website and in Sri Lankan newspapers.

But it is not just contacts that make a difference. Sometimes society membership provides insight into how institutions operate. Patrick Tuijp, treasurer of the European Council of Doctoral Candidates and Junior Researchers (Eurodoc) in Brussels, says that his role on the council — a federation of 34 national organizations — has taught him about European funding schemes. He has learned in particular about Horizon 2020, the European Union's main research-funding mechanism for 2014–20, and the Marie Curie Actions mobility research grants. He understands how they function, who is eligible and how to improve the likelihood of winning grants — which in turn helps him to advise other doctoral students and informs his own decisions. “I have a clearer grasp on what’s going to happen in five years,” says Tuijp, an assistant professor of finance at the University of Amsterdam and a PhD student in financial economics at Tilburg University in the Netherlands.

On a more altruistic level, says Tracey, the NPA offers emotional and strategic support to its members, and she joined the association to be part of that. “My driving reason for wanting to be involved is to give back,” she says. She has helped US funding agencies to understand the importance of mentoring for postdocs; as a result of the NPA's work, the NSF now requires its grant applicants to set out a plan for how they will mentor their postdocs. Tracey has also worked to increase the US National Institutes of Health's Ruth L. Kirschstein National Research Service Award postdoc stipend, on which many universities base their postdoc salaries. “I got a lot of satisfaction out of that,” she says.

PARTICIPATION

Trouble maintaining members

It is not easy to keep an association or society powering along when its membership is transient and short-term. A postdoc or contract researcher's highest priority is their next job, which often means that they are not committed to membership or leadership of a society, says Nicola Woodward, a scientist at the Institute of Food Research in Norwich, UK, and co-chair of the committee of the UK Research Staff Association (UKRSA) in Cambridge.

She says that some UKRSA members avoid taking on projects that might continue after their current posts end. As young researchers move on to a new postdoc or a permanent job, it can be difficult to recruit more members or fill administrative posts. When administrative posts are left vacant, it can adversely affect member services and events. And a smaller membership means fewer new colleagues to meet, network with and exchange ideas with.

The US National Postdoctoral Association (NPA) in Washington DC also struggles with recruitment and service hurdles, says Ian Brooks, its international officer and director of the office of biomedical informatics at the University

of Tennessee Health Science Center in Memphis. He thinks that some prospective members, knowing that postdocs receive little respect at many institutions, may refrain from embracing the label and joining the association. “That’s the biggest hurdle the NPA faces,” he says.

Similar obstacles face the International Consortium of Research Staff Associations (ICoRSA), launched last year as an umbrella organization for groups including the NPA and the UKRSA. Despite an international pool of potential members, ICoRSA, which is based in Cork, Ireland, is having trouble recruiting, says chair Gordon Dalton, a senior research fellow in ocean energy economics at University College Cork.

Dalton, who was active in his university and national research-staff associations before the ICoRSA formed, knew that he wanted to serve but says that he can do so only because he has a seven-year renewable research contract. Many researchers have too many other duties and feel as if their principal investigators (PIs) are looking over their shoulders. “Some people attend meetings,” he says, “and ask that their PIs not be told.” **K.K.**

DEVELOPING CONNECTIONS

Membership of global scientific associations and academies can be especially useful for scientists in developing countries. “It’s provided opportunities to establish collaborations with scientists I would not meet in my regular line of work,” says GYA member Patrick Arthur, a biochemist at the University of Ghana in Accra. He is working with an analytical chemist in Egypt on the effects of aluminium leaching into food from cookware, and with a group in the Netherlands to seek medically useful compounds in wild mushrooms. International collaborations increase researchers’ chances of getting grants, he notes, and lead to improved visibility and more invitations to present at prestigious conferences. These, in turn, lead to further collaborations and funding opportunities (see ‘Joining up’).

Thadhani agrees. “You are made aware of various scholarships, awards, conferences and workshops, in addition to meeting people you can collaborate with,” she says. Global-association membership “helps in bridging

NETWORKING

Joining up

Most professional associations for early-career researchers require only that members fit certain criteria, such as being a postdoc in a particular country. Others have stricter rules.

- **Global Young Academy, Berlin:** Members are chosen for the excellence of their science and their commitment to solving global problems. Prospective members must apply with a letter of support from their national academy, an equivalent body, their employer, their institution or another professional.

- **World Association of Young Scientists:** Membership is open to any early-career scientist who agrees with the association's goals, including promoting excellence and helping young scientists in their careers.

- **US National Postdoctoral Association, Washington DC:** Open, with varying membership fees, to any graduate student or postdoctoral researcher from any nation who endorses the association's mission of supporting the postdoctoral experience.

- **UK Research Staff Association, Cambridge:** Any UK early-career researcher can join to interact online, participate in activities or get involved in the advisory group.

- **International Consortium of Research Staff Associations, Cork, Ireland:** Membership is open to early-career researchers who belong to a research-staff association in a member nation.

- **Eurodoc, Brussels:** Members must belong to a research-staff association that represents doctoral candidates and/or junior researchers in a European Union or Council of Europe member state. If their country does not have such a group, researchers may be able to join with observer status. **K.K.**

the gap between developing and developed countries”.

“It’s all about extending your networks and building new networks,” says Nicola Woodward, co-chair of the committee of the UK Research Staff Association in Cambridge. “The whole focus is to encourage people to expand.” ■

Karen Kaplan is associate editor of *Nature Careers*.

COLUMN

A good investment

Success involves acknowledging past accomplishments as well as looking ahead to future value, says **Yoshimi Rii**.

When I became the inaugural recipient of a research fellowship this year, my department commemorated the occasion with a ceremony for which they asked me to prepare a ten-minute talk. I was to thank the foundation that funded the fellowship, and describe my research. I have given many talks in my life, but I found myself stumped as to what I should focus on.

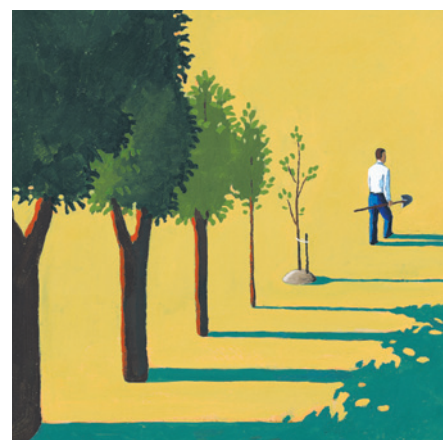
“Keep it pretty simple on the science, because they want to know who you are,” the foundation director advised me. I can easily talk for ten minutes about the role of phytoplankton in nutrient cycling, even slipping in some poop jokes, but this time the talk had to focus on me. How was I supposed to reassure the foundation that I deserved its investment while still sounding humble?

I thought about what ‘investment’ really means. Investing involves expectation of future gain. To invest is to believe in potential, and a good investment is gauged by the end result.

The thought of being a good investment riddled me with anxiety. With financial freedom came an ocean of expectations. Was I really worthy of this award, and why? I suddenly feared for my future, wondering whether I would make it in the scientific world and uphold the legacy of oceanographic research. Would the foundation still be proud of me if I did not end up pursuing a postdoc and the conventional route to academia? Would I still be able to call myself a scientist? I am 34. I got married in August, and starting a family was on my mind — and still is. Was it wrong to consider getting pregnant while on this fellowship?

I grew obsessed with what I would become, but it was my present self that had won the fellowship. As I drew up the outline for my talk, I tried to focus on my PhD journey. I thought about being in the right place at the right time, and about the collaborations and the help of many wonderful people that got me here. My ten-minute talk was starting to sound like a list of acknowledgements at the Oscars.

That night, I watched Sheryl Sandberg, chief operating officer of Facebook and author of *Lean In: Women, Work, and the Will to Lead* (Knopf, 2013), in an interview on the television news programme *60 Minutes*. “Women attribute their success to working hard, luck and help from other people,” she said. “Men will attribute that same success to their own core skills.” Sandberg insisted that the reason



JONATHAN EVANS/GETTY

there are fewer women than men in top leadership roles is that women hold themselves back.

I listened with fascination. I did not see myself as someone who leaned back, but here I was, attributing my fellowship to everyone else and completely anxious about my future job and an imaginary baby.

Empowered by Sandberg's words, I re-evaluated how I should be leaning in with my speech. To move others, I needed to draw on my own inspirations and reflect on why I study the ocean. Most of the time, I am too exhausted, too cynical and too concerned with minute details to take a step back and acknowledge that I am here because I put myself here. But at some point between the sleepless hours at sea and in the lab, I became a person worth investing in. So in my speech, I told the foundation members about my first research cruise 11 years ago, when I threw up for five days straight. I told them how I almost quit grad school when my PhD adviser moved across the continent, and how grateful and honoured I felt to be standing in front of them, supported by my friends and colleagues. Persistence, I felt, is something anyone can relate to — and find worthy of investment.

Pressures and anxieties will always be there. But I learned how important it is, as graduate students and postdocs — and eventually as professors, educators, industry managers and whatever else we hope to be — to focus not only on what we will become, but also on who we are now. ■

Yoshimi Rii is a graduate student in microbial oceanography at the University of Hawaii at Manoa.

HOW CHERRY COKE SAVED MY LIFE

A lucky break.

BY DAWN BONANNO

“A meetings didn’t work for me,” I told the little alien. I didn’t know what to do with him. He kind of looked like a poodle, so it seemed like I was sitting on my front stoop talking to a dog I didn’t own. As if the hawk-eye neighbours didn’t think I was strange enough, living in this big house all by my lonesome. Well, what used to be my house. At least the pond was still intact, pretty as it reflected the afternoon sun in purple ripples.

“A . . . A?” His puppy-dog eyes crossed.

“A place where drunks go to cure themselves. Works for millions of Americans. Just not of Harold. It was the crowds, you see. Even the smaller meetings were too crowded for me. Some guy offered to mentor me solo, but then I’d lose the anonymous part of the deal. The guy would know me. People would hear about the screwball who was so far gone even AA couldn’t help him.”

Poodle guy nodded, his ear curls bouncing. “My translator is functioning properly now. Allow me to introduce —”

I grabbed his snout. Not hard, just enough to make him shut up. “Anonymity, dude. Work with me here.” Not giving him a chance to realize I’d already introduced myself, I went on. “So that was around when Jenna left me, and I was really in hot water. I don’t cook. Sure as hell can’t clean. She took care of me and the kids, but after I got fired, she was done. I almost died without her.”

“Is that when the Cherry Coke saved your life?”

“Not yet, but it is when I started on the Cherry Coke. Had to transfer my dependence onto something that wouldn’t get me fired again or permanently divorced or smash my wheels. Figured the Coke was a good idea as I’d already done in my liver with all that drinking. As long as I drink extra water and run a bit, my kidneys’ll hold up. The running did me some serious good. It lost me thirty pounds.”

I felt like scratching behind his ears, but he’d probably take offence to that, being new to Earth and all, and not knowing our

relationship with the resident four-leggers. Have to give him credit for just looking at me all strange.

“Thirty pounds.” Poodle guy cocked his head to the side then nodded. “Removing 15% of your body mass would seriously relieve the stress on your internal organs. Is that how Cherry Coke saved your life?”

“Nah. The running was good, but it wasn’t life-changing or nothing.” What was he, a doctor or something? If that was the case, he wouldn’t be much help with my house. “See, I still want Jenna back. I called her today, told her about all I done, invited her over. Damn but she said yes! I started shaking all over, and that’s when I realized I needed a Coke. Crappy timing though, I was all out.”

The poodle guy nodded, his ears flopping, as if making some big discovery from my troubles. “Out of inventory?”

I snorted. “I put a dent in the A&P’s inventory, filled up my trunk. Didn’t even stop on the way back for gas, which my wheels needed to get me back here. So I was late for Jenna. Saw her

driving back and caught some nasty words, but she never did slow down. Don’t think she’ll be back.”

“Is that how Cherry Coke saved your life?”

“Nah.” I grinned now and this time I pet him on the head. Probably confused him more than it ticked him off. I took a long swig of my warm Cherry Coke, finishing off the can. “See, the Cherry Coke? It got me out of the house and stuck on those back roads when you lost control of your damned saucer —”

“Spacepod.”

“I don’t care what you call it, it trashed my house. Look at it! How the hell am I supposed to live in a pile of smashed timber and space metal? That’s what I was thinking as I walked up, soda case in hand. Yep. If it weren’t for the Cherry Coke, I’d have been home, pacing my living room when your saucer crashed down onto it. So you see, my furry friend, that’s how Cherry Coke saved my life.”

Poodle guy was silent as he stared at the wreckage of my empty home. Not that I cared about the house itself. Like I said, can’t clean. If Jenna had actually gone in, she’d have left me again on account of being a slob. At least I had a dog now. Sort of.

“Welcome to Earth, friend,” I said.

“Oh no,” he said and stared wide-eyed at the wreckage.

“What, dude?”

“Wrong planet.” He whimpered now.

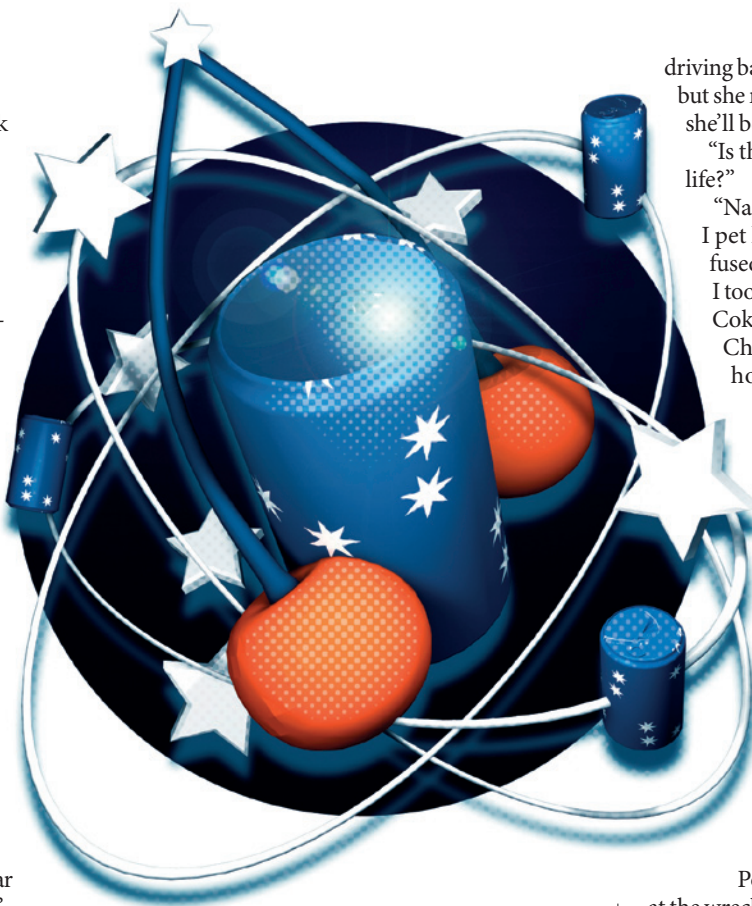
“That sucks. Will your friends come rescue you?”

“I was supposed to rescue them.”

The debris that had been my house shifted, the timber and metal collapsing in on itself. Homeless is homeless, it didn’t matter what that mess did now.

“Shit happens, friend,” I said, deciding to keep him, and opened him up a warm can of Cherry Coke. “Best you can do is live in the moment. You never know what might come of it.” ■

Dawn Bonanno suffers from an obsession with pens, paper and fixing things, so it only makes sense that she writes stories. Rumours of her Cherry Coke addiction have been documented at www.dmbonanno.com.



➔ NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

Controversy about ultrahard nanotwinned cBN

ARISING FROM Y. Tian *et al.* *Nature* **493**, 385–388 (2013)

Tian *et al.*¹ report synthesis of “nanotwinned” cubic boron nitride (“nt-cBN”). These authors claim that its unprecedented Vickers hardness of 108 GPa is due to nanotwinning, and that hardening of cBN is continuous with decreasing twin thickness down to the smallest size investigated, in contrast to the expected reverse Hall–Petch effect. We demonstrate here that it has been known for tens of years that “nt-cBN” (refs 2–4) hardens owing to a complex of phenomena associated with its microstructure and defects; we also consider that Tian *et al.*¹ provide no proof that the hardness of “nt-cBN” is larger than 85 GPa, which is the previously reported maximum for nanostructured boron nitride^{2,4}. Thus the claim of continuous hardening down to a few nanometres twin thickness is, in our opinion, unjustified. There is a Reply to this Brief Communication Arising by Tian, Y. *et al.* *Nature* **502**, <http://dx.doi.org/10.1038/nature12621> (2013).

The microstructure (small grain sizes starting from submicrometre values) and structural defects (twinning, stacking faults and dislocations) of a material strongly affect its X-ray diffraction and Raman scattering properties. The effects are manifested in diffraction peak broadening, asymmetry, *hkl*-dependent intensity variations, and the appearance of extraordinary reflections characteristic of nanotwinned structures^{4–6}. We note that the X-ray diffraction patterns and Raman spectra of bulk “nt-cBN” presented by Tian *et al.*¹ in their Supplementary Figure 2 strictly disagree with the declared samples’ nanostructure. They display sharp peaks characteristic of polycrystalline (well-crystallized) materials with no features due to nanocrystallinity, nanotwinning or stacking faults. The Raman spectra provide no indication of the quantum confinement effect (supposed to be responsible for “nt-cBN” hardening), which has been observed in the spectra of various nanomaterials^{4,7} including boron nitride⁴ with crystallite sizes a few times larger than those reported by Tian *et al.*¹ Thus, the data shown in Supplementary Figure 2 of ref. 1 are not related to the “nt-cBN” described in the main part of the text (Figs 1 and 2)¹. The reason could be inhomogeneity of samples containing both a well-crystallized component, giving a major

contribution in X-ray diffraction patterns and Raman spectra, and a nanocrystalline component (investigated by transmission electron microscopy, TEM). Thus, only the TEM data¹ are related to the nanomaterial, and we consider only these data below.

For materials with a diamond-like structure, lamellar {111} twins and stacking faults parallel to the (111) plane mean the alternation of larger or smaller blocks of diamond-like close-packed layers stacked along the {111} direction^{8,9}, either in a ‘cubic’ (*c*) or ‘hexagonal’ (*h*) type of packing. These two types of packing correspond to the two structures of *sp*³-bonded dense BN polymorphs: namely 3C cBN (cubic sphalerite-type structure, only *c* layers) and 2H wBN (hexagonal wurtzite-type structure, only *h* layers) (Fig. 1). Such an arrangement was described in ref. 4 as alternation “of blocks with 2-layered wurtzite and 3-layered sphalerite structure... at the level of a few to a dozen of nanometers”; it was used to describe the fine structure of the 20–70 nm sized particles of nano-BN studied in ref. 4 and called there “aggregated boron nitride nanocomposite”, ABNNC (Fig. 2). Tian *et al.*¹ call the building blocks “nanotwins” neglecting stacking faults (as large as or larger than “twins”, see Figure 2b in ref. 1) and grain boundaries, although it is exactly a combination of all these structural features that is responsible for hardening of the material^{2–4,9}.

We consider that the results of hardness measurements in ref. 1 are unconvincing. Tian *et al.*¹ write that the Vickers hardness H_V “of the nt-cBN bulk decreases from ~196 GPa at 0.2 N to its asymptotic value, 108 GPa, beyond 3 N”. First, we maintain that a hardness of 196 GPa not only has no physical meaning^{10–12}, but could not be measured in principle (an indentation giving $H_V \approx 196$ GPa would have to have a diagonal length of about 1.3 μm , comparable with or even smaller than the size of mechanical polishing defects of the sample surface visible in the left inset of Figure 3 in ref. 1). Second, the indentation shown in this inset¹ displays remarkably short cracks made under the load of 19.6 N and has an average diagonal length of ~21 μm , which gives $H_V \approx 82$ GPa: this value is significantly lower than claimed maximum of 108 GPa (ref. 1). It suggests that the flattening of the “ H_V versus load” curve between 3 and 7 N (in Figure 3 of ref. 1) is artificial; we note that much higher loads are required for saturation of population growth of microcracks in a brittle material¹⁰ to report a correct hardness value. In fact, $H_V \approx 82$ GPa for “nt-cBN” is similar within the experimental error to the value reported for ABNNC (85(5) GPa)⁴.

The Knoop indenter is more appropriate for measuring the hardness of superhard materials^{13,14}. The reported Knoop hardness of “nt-cBN” ($H_K = 77.7 \pm 3.8$ GPa)¹ perfectly matches that of the material^{2,4}

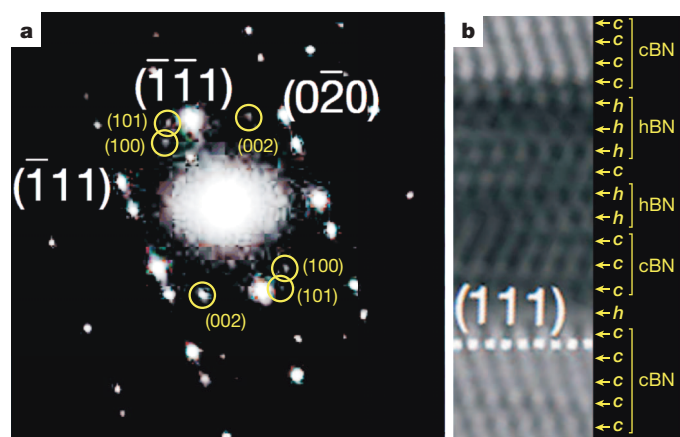


Figure 1 | TEM data for “nt-cBN”. **a**, Electron diffraction pattern of “nt-cBN” (see inset in Figure 2b of ref. 1) displaying unidentified diffraction spots (highlighted by circles) which can originate from different grains of wBN and be assigned as designated (yellow numbers). These data suggest that the material, declared in Tian *et al.*¹ as pure cBN, indeed contains substantially large wBN-type domains. **b**, Part of the high-resolution TEM image of “nt-cBN” (Figure 2b in ref. 1), where the layers corresponding to cBN are marked with ‘c’, and those corresponding to wBN with ‘h’.

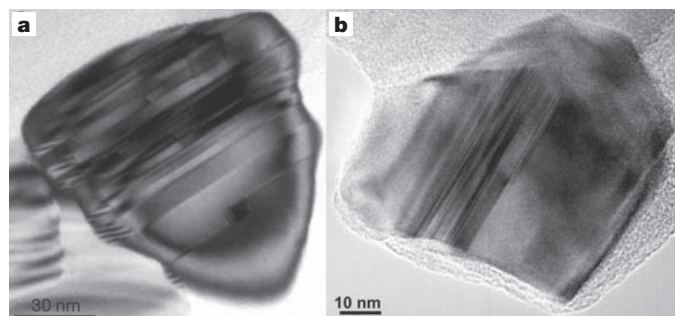


Figure 2 | Bright-field TEM images. **a**, “nt-cBN” (Figure 2a in ref. 1) and **b**, ABNNC (ref. 4), demonstrating similarity of the microstructure of these materials.

obtained from pyrolytic BN precursors ($H_K \approx 78$ GPa)². We note that the conclusion of Tian *et al.* — that “nt-cBN” does not show the reverse Hall–Petch effect (and its hardening is continuous)—is based on the only point (“3.8 nm – 108 GPa”) added by Tian *et al.*¹ to the data set reported in ref. 4 for ABNNC. If the H_V of “nt-cBN”¹ is in fact a maximum of 82 GPa, as we argue above, the nanocrystalline sp^3 -bonded BN seems to not be an exception and indeed, like other materials, shows the reverse Hall–Petch effect.

Natalia Dubrovinskaia¹ & Leonid Dubrovinsky²

¹Material Physics and Technology at Extreme Conditions, Laboratory of Crystallography, University of Bayreuth, 95440 Bayreuth, Germany. email: natalia.dubrovinskaia@uni-bayreuth.de

²Bayerisches Geoinstitut, University of Bayreuth, 95440 Bayreuth, Germany.

Received 19 February; accepted 20 August 2013.

1. Tian, Y. *et al.* Ultrahard nanotwinned cubic boron nitride. *Nature* **493**, 385–388 (2013).
2. Corrigan, F. R. & Bundy, F. P. Direct transitions among the allotropic forms of boron nitride at high pressures and temperatures. *J. Chem. Phys.* **63**, 3812–3820 (1975).
3. Horiuchi, S., He, L.-L., Huang, J., Taniguchi, T. & Akaishi, M. Development of superhard materials using HRTEM. *J. Surf. Anal.* **3**, 197–202 (1997).

4. Dubrovinskaia, N. *et al.* Superhard nanocomposite of dense polymorphs of boron nitride: noncarbon material has reached diamond hardness. *Appl. Phys. Lett.* **90**, 101912 (2007).
5. Solozhenko, V. L., Kurakevych, O. O. & Le Godec, Y. Creation of nanostructures by extreme conditions: high-pressure synthesis of ultrahard nanocrystalline cubic boron nitride. *Adv. Mater.* **24**, 1540–1544 (2012).
6. Rao, W.-F. & Wang, YuU Diffraction theory of nanotwin superlattices with low symmetry phase: adaptive diffraction of imperfect nanotwin superlattices. *Phil. Mag.* **90**, 197–217 (2010).
7. Bersani, D., Lott, P. P. & Ding, X.-Z. Phonon confinement effects in the Raman scattering by TiO₂ nanocrystals. *Appl. Phys. Lett.* **72**, 73–75 (1998).
8. Wells, A. F. *Structural Inorganic Chemistry* (Oxford Univ. Press, 1975).
9. Dubrovinskaia, N. & Dubrovinsky, L. in *High-Pressure Crystallography: From Fundamental Phenomena to Technological Applications* (eds Boldyreva, E. & Dera, P.) 419–434 (Springer Science and Business Media, 2010).
10. Brazhkin, V. *et al.* What does ‘harder than diamond’ mean? *Nature Mater.* **3**, 576–577 (2004).
11. Eremets, M. I. *et al.* The strength of diamond. *Appl. Phys. Lett.* **87**, 141902 (2005).
12. Dubrovinsky, L., Dubrovinskaia, N., Prakapenka, V. B. & Abakumov, A. M. Implementation of micro-ball nanodiamond anvils for high-pressure studies above 6 Mbar. *Nature Commun.* **3**, 1163 (2012).
13. Sumiya, H. & Irifune, T. Indentation hardness of nano-polycrystalline diamond prepared from graphite by direct conversion. *Diamond Relat. Mater.* **13**, 1771–1776 (2004).
14. Sumiya, H. Super-hard diamond indenter prepared from high purity synthetic diamond crystal. *Rev. Sci. Instrum.* **76**, 026112 (2005).

Author Contributions N.D. and L.D. contributed equally to this paper.

Competing Financial Interests Declared none.

doi:10.1038/nature12620

Tian *et al.* reply

REPLYING TO N. Dubrovinskaia & L. Dubrovinsky *Nature* **502**, <http://dx.doi.org/10.1038/nature12620> (2013)

Dubrovinskaia and Dubrovinsky¹ question the sub-grain structure and ultrahigh hardness of our nanotwinned cubic boron nitride (nt-cBN; ref. 2), and make the following assertions: (1) hardening related to nanotwinning in cBN has been known for tens of years, (2) our X-ray diffraction (XRD) and Raman results do not show the signatures of nanotwinning, (3) there are large wurtzite-structure boron nitride (wBN) domains present in our samples, and (4) our hardness values are unconvincing. We argue below that all these claims are incorrect.

Assertion (1). Although there are indeed examples of nanostructured cBN in the literature, single-phase cBN with ubiquitous nanotwinned submicrostructure has never been reported previously². In fact, none of the earlier studies of polycrystalline cBN involved a hardening mechanism based on nanotwinning^{3–5}. Nanotwinning was not mentioned in refs 3 or 4 or references cited therein, and neither was nanotwinning-induced hardening.

Assertion (2). XRD and Raman data from the same sample described in the main text of ref. 2 are presented here in Fig. 1a and b, along with those from previous ABNNC sample³, on which the main claims of Dubrovinskaia and Dubrovinsky¹ are based. The differences between nt-cBN (black curves) and ABNNC (red curves) are striking. In our nt-cBN samples, adjacent {111} twins share a coherent boundary. The atoms within the twin boundary are arranged in the same manner as those inside twin domains and do not exhibit lattice distortion. Such coherent twin boundaries do not contribute remarkably to the broadening of XRD or Raman peaks. The grain size in nt-cBN (30–150 nm) is larger than that of ABNNC (14 nm), leading to sharper XRD and Raman peaks in our samples. Also, the quantum confinement effect proposed in ref. 2 is different from the phonon confinement effect referred to in ref. 1, and is not detectable by Raman measurement.

Assertion (3). Transmission electron microscopy (TEM) observations indicate the ubiquity of nanotwins inside every nt-cBN nanograin, as

well as the microstructural homogeneity of our samples. Dubrovinskaia and Dubrovinsky¹ attributed the faint “unidentified diffraction spots” in the background of Figure 1a of ref. 1 to “different grains of wBN”, which is rather arbitrary. Both XRD and Raman measurements exclude the existence of wBN in our samples. It is more reasonable to relate these spots to different cBN grains. For high resolution TEM (HRTEM) imaging, Dubrovinskaia and Dubrovinsky¹ selected from Figure 2b of ref. 2 an area across twin boundaries with a width of only several atoms, and interpreted the stacking defects as atomic layers of a wBN phase (Figure 1b of ref. 1). We argue that this is misleading. Our HRTEM analyses clearly show that these defects are localized and span only a few atoms. The existence of wBN phase is thus ruled out.

Assertion (4). The indenter size effect results in higher hardness values in smaller indentations owing to a greater strain gradient⁶, thus a reliable hardness should be determined from the asymptotic-hardness region of a well-controlled indentation process. The asymptotic hardness of nt-cBN was determined in the same way as that reported for ABNNC (ref. 3). Figure 1c compares the hardness–load curves for the two materials. The superior hardness of nt-cBN over ABNNC is clear. The continuous hardening down to a few nanometres (Figure 4 of ref. 2) is therefore justified for nt-cBN, contrasting sharply with the reversal in hardness as reported for nanograined cBN and ABNNC (ref. 3). The Vickers hardness values obtained from loads over 10 N are not valid because there are large cracks formed around the indentation⁷. In addition, the Knoop hardness reported in ref. 4 was averaged at 69 GPa (78 GPa as the maximum), which is lower than the 77.7 GPa of our nt-cBN (ref. 2).

Can hardness be reliably measured for materials, if any, harder than natural diamond? This question has been vexing researchers for a long time^{7,8}. Our answer is yes. Indentation hardness is determined by load divided by the projected area of a permanently formed indentation⁷. In

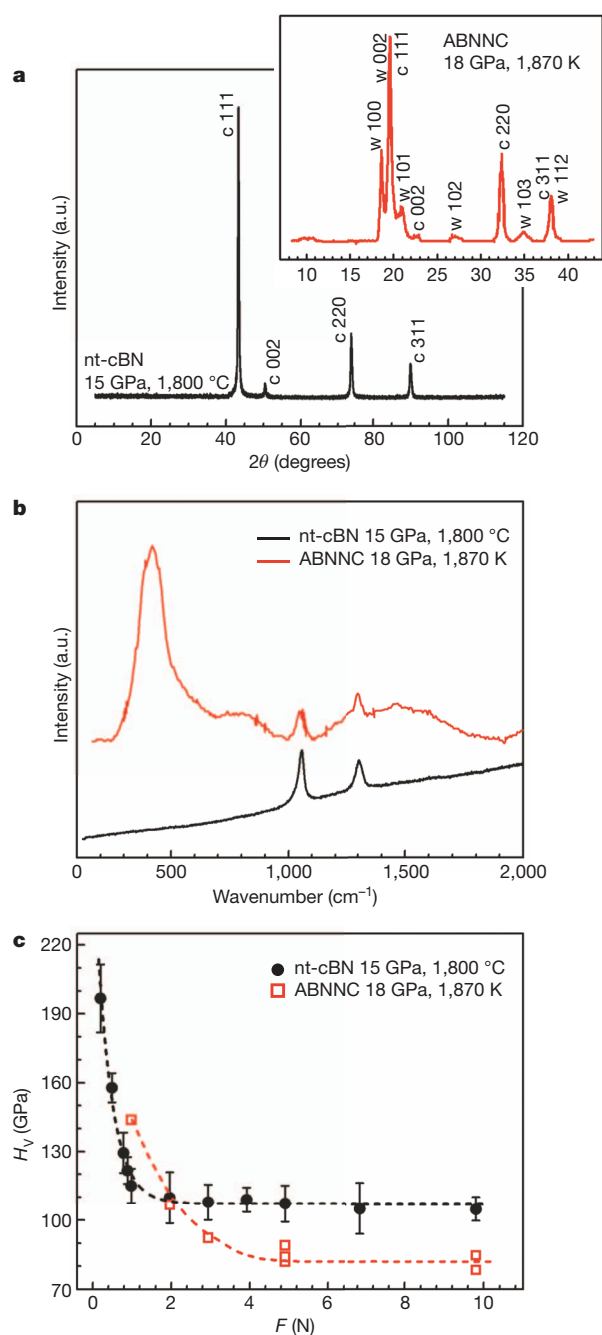


Figure 1 | Materials properties for representative nt-cBN (ref. 2) and ABNNC (ref. 3) samples. a, XRD patterns; b, Raman spectra; and c, load-dependent Vickers hardness. Error bars on nt-cBN data in c indicate s.d. ($n = 5$).

measurement, stress states are different in the indenter and the sample's tested zone: the tip of the diamond indenter is subjected to a compressive stress field, and the sample undergoes plastic shear deformation around the indenter. The compressive strengths of diamond are 223 GPa in the weakest $\langle 100 \rangle$ direction, and about 470 GPa along $\langle 110 \rangle$ and $\langle 111 \rangle$ (ref. 9), whereas the shear strengths of diamond and cBN are 93 GPa (ref. 10) and 65 GPa (ref. 11), respectively. Indentation hardness can be measured reliably as long as the shear strength of the sample is smaller than the compressive strength of the indenter diamond. This requirement is satisfied even if the measured sample is harder than natural diamond, for example, 170 GPa for annealed CVD diamond¹². Therefore, the hardness exceeding that of diamond does have physical meaning, even for values as high as several hundred gigapascals.

Yongjun Tian¹, Bo Xu¹, Dongli Yu¹, Yanming Ma², Yanbin Wang³, Yingbing Jiang⁴, Wentao Hu¹, Chengchun Tang⁵, Yufei Gao¹, Kun Luo¹, Zhisheng Zhao¹, Li-Min Wang¹, Bin Wen¹, Julong He¹ & Zhongyuan Liu¹

¹State Key Laboratory of Metastable Materials Science and Technology, Yanshan University, Qinhuangdao 066004, China.

email: fhcl@ysu.edu.cn

²State Key Laboratory for Superhard Materials, Jilin University, Changchun 130012, China.

³Center for Advanced Radiation Sources, University of Chicago, Chicago, Illinois 60439, USA.

⁴TEM Laboratory, University of New Mexico, Albuquerque, New Mexico 87131, USA.

⁵School of Material Science and Engineering, Hebei University of Technology, Tianjin 300130, China.

1. Dubrovinskaya, N. & Dubrovinsky, L. Controversy about ultrahard nanotwinned cBN. *Nature* **502**, <http://dx.doi.org/10.1038/nature12620> (2013).
2. Tian, Y. *et al.* Ultrahard nanotwinned cubic boron nitride. *Nature* **493**, 385–388 (2013).
3. Dubrovinskaya, N. *et al.* Superhard nanocomposite of dense polymorphs of boron nitride: noncarbon material has reached diamond hardness. *Appl. Phys. Lett.* **90**, 101912 (2007).
4. Corrigan, F. R. & Bundy, F. P. Direct transitions among the allotropic forms of boron nitride at high pressures and temperatures. *J. Chem. Phys.* **63**, 3812–3820 (1975).
5. Horiuchi, S., He, L.-L., Huang, J., Taniguchi, T. & Akaishi, M. Development of superhard materials using HRTEM. *J. Surf. Anal.* **3**, 197–202 (1997).
6. Nix, W. D. & Gao, H. Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J. Mech. Phys. Solids* **46**, 411–425 (1998).
7. Chaudhri, M. M. & Lim, Y. Y. Harder than diamond? Just fiction. *Nature Mater.* **4**, 4 (2005).
8. Brazhkin, V. *et al.* What does 'harder than diamond' mean? *Nature Mater.* **3**, 576–577 (2004).
9. Luo, X. *et al.* Compressive strength of diamond from first-principles calculation. *J. Phys. Chem. C* **114**, 17851–17853 (2010).
10. Roundy, D. & Cohen, M. L. Ideal strength of diamond, Si, and Ge. *Phys. Rev. B* **64**, 212103 (2001).
11. Pan, Z., Sun, H., Zhang, Y. & Chen, C. Harder than diamond: superior indentation strength of wurtzite BN and lonsdaleite. *Phys. Rev. Lett.* **102**, 055503 (2009).
12. Yan, C. S. *et al.* Ultrahard diamond single crystals from chemical vapor deposition. *Phys. Status Solidi A* **201**, R25–R27 (2004).

doi:10.1038/nature12621